Micromachined Directional Hydrophones

Final Report

Thomas B. Gabrielson
Applied Research Laboratory
The Pennsylvania State University
P.O. Box 30
State College, PA 16804
tbg3@psu.edu

Contract N00014-97-1-0062
Program Officer: Dr. Roy C. Elswick
ONR 321SS
Office of Naval Research
Ballston Centre Tower One
800 North Quincy Street
Arlington, VA 22217-5660

7 August 2000

ABSTRACT

Regional actions in littoral waters will likely take the form of rapidly developing conflicts involving multiple surface-ship, mine, and submarine threats. The mission effectiveness of various Navy assets would be greatly enhanced with a deployable multi-threat sensor system for underwater warfare. One critical component of such a system is a miniature, directional hydrophone (or acoustic velocity sensor). While useful in platform systems as well, miniaturization of velocity hydrophones has the greatest payback in deployable systems – either sonobuoy systems or longer-life deployable systems. Current directional sonobuoys employ accelerometer-based velocity sensors using piezoelectric materials for transduction. However, miniaturization favors other forms of transduction. One form that is particularly suited to microfabrication is the differential-capacitance accelerometer. The differential-capacitance configuration is capable of good long-term stability and exceptional low-frequency noise performance. The effective sensor impedance is determined not by the signal frequency but by the AC drive frequency. The differential-capacitance accelerometer does not have an inherent roll-off in response at low frequency so these devices can be used in self-orienting three-axis sensors for deployable systems. This report documents an investigation into the fundamental limits associated with miniaturization of velocity hydrophones and the development of the differential-capacitive sensor as a directional hydrophone element.

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

20000905 105

DTIC QUALITY INSPECTED 4

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	7 August 2000	Final 01 October 1996 - 30 June 2000

4. TITLE AND SUBTITLE
Micromachined Directional Hydrophones

5. FUNDING NUMBERS
G: N00014-97-1-0062
PR: 99PR01497-00

6. AUTHORS
Thomas B. Gabrielson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Pennsylvania State University
Office of Sponsored Programs
110 Technology Center Building
University Park, PA 16802-7000

8. PERFORMING ORGANIZATION REPORT NUMBER
FD-321-00-01

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research (ONR 252: Diane Gales)
Balston Centre Tower One
800 North Quincy Street
Arlington, VA 22217-5560

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution is Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Regional actions in littoral waters will likely take the form of rapidly developing conflicts involving multiple surface-ship, mine, and submarine threats. The mission effectiveness of various Navy assets would be greatly enhanced with a deployable multi-threat sensor system for underwater warfare. One critical component of such a system is a miniature, directional hydrophone (or acoustic velocity sensor). While useful in platform systems as well, miniaturization of velocity hydrophones has the greatest payback in deployable systems – either sonobuoy systems or longer-life deployable systems. Current directional sonobuoys employ accelerometer-based velocity sensors using piezoelectric materials for transduction. However, miniaturization favors other forms of transduction. One form that is particularly suited to microfabrication is the differential-capacitance accelerometer. The differential-capacitance configuration is capable of good long-term stability and exceptional low-frequency noise performance. The effective sensor impedance is determined not by the signal frequency but by the AC drive frequency. The differential-capacitance accelerometer does not have an inherent roll-off in response at low frequency so these devices can be used in self-orienting three-axis sensors for deployable systems.

14. SUBJECT TERMS
Acoustics, USW, Velocity Sensor, Directional Hydrophone, Accelerometer, Differential Capacitance

15. NUMBER OF PAGES
82

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
U	U	U	UL

Micromachined Directional Hydrophones

Table of Contents

Executive Summary ii
Lessons Learned: Microfabrication iii
Lessons Learned: Velocity Sensors v

Chapter 1: Introduction 1-1

Chapter 2: Performance Limits: Sensor Self-Noise 2-1
Basic Thermal Fluctuations 2-1
Noise from Molecular Collisions 2-6

Chapter 3: Performance Limits: Sensor/Amplifier Interaction 3-1
Fundamentals 3-1
Transistor Noise 3-3
Interaction between Sensors and Amplifiers 3-7

Chapter 4: Structural Limits in Microfabrication 4-1
Basic Scaling 4-1
Microfabrication 4-2
Micro-Arrays of Sensors 4-3
Residual Stress Management 4-4
Fluid Dynamics 4-10
Electrostatic Forces 4-14
Three-Dimensional Effects 4-16

Chapter 5: The Differential-Capacitance Hydrophone Element 5-1
Capacitive Sensing 5-1
Basic Configuration 5-4
System Noise Analysis 5-5
Evaluation 5-8
Multiplexing AC-Drive Sensors 5-12

Chapter 6: Velocity Sensor Suspension Dynamics 6-1
Basic Velocity-Sensor Dynamics 6-1
Influence of the Suspension 6-5
Types of Suspensions 6-7
Suspension Anisotropy 6-8
Impedance-Matching Material 6-11
Arrays of Sensors in Compliant Layers 6-12

Appendix A: Publications, Presentations, and Patents A-1

Executive Summary

The long-term goal of this investigation was to develop a high-performance, miniature, directional hydrophone to either replace existing omnidirectional sensors or to greatly increase the number of directional sensors in deployable surveillance and tactical ASW and USW systems. Increased directionality yields improved signal detection, improved noise discrimination, and improved classification ability.

The technical objective was to develop a low frequency, miniature, directional hydrophone element by exploiting emerging silicon microfabrication technology (MicroElectroMechanical Systems or MEMS). The design objectives were:

- a packaged two-axis sensor with a volume of 8 cm^3
- self noise less than or equal to the equivalent of Sea State Zero - Light Shipping
- useable frequency range from 5 Hz to 1000 Hz
- linearity in response over at least 70 dB
- hydrostatic pressure tolerance to 600 m (2000 ft) depth.

The volume requirement represents better than a 50:1 volume reduction over directional hydrophones used in current sonobuoys. The self-noise requirement is equivalent to acoustic accelerations in the 10 to 100 nano-g (1 Hz band) range.

The differential-capacitance accelerometer has the theoretical potential for meeting the self-noise requirement in a package of acceptable size. In contrast to more sophisticated microfabricated devices, commercial off-the-shelf (COTS) versions of the differential-capacitance accelerometer are available from several manufacturers. One such commercial device was adapted for use as a directional hydrophone element in this investigation. Most aspects of the desired performance were demonstrated; however, the proof mass of the commercial device was too small to produce the required self-noise level. While the measured self-noise was several hundred nano-g per root hertz, there is a plausible path for reduction of the self-noise to the required level. The critical change is to enlarge the device and increase the mass of the proof mass. Since the COTS device is much smaller than required for the hydrophone application, this change can be accomplished.

Many aspects of this investigation highlighted practical problems with microfabrication of high-performance sensors and with acoustic particle velocity sensing. Consequently, the following two sections of "Lessons Learned" are included as a guide to other workers in the field. The lessons have certainly been biased by the particular application; however, there is enough generality in many of the principles that they are applicable beyond the design of low-noise accelerometers as velocity hydrophones.

Lessons Learned: Microfabrication

In order to increase the chance for highly successful programs, it is important to understand why MEMS development efforts fail. The primary reasons are:

- Poor mechanical design
- Poor electrical design
- Incompatible processes
- Residual stress
- Disregard for packaging

Microfabrication is not a panacea. Most descriptions of MEMS dwell on the positive aspects and the potential; however, there are many negatives including high uncertainty in eventual cost; high self-noise; device variability; and a lack of standard, reliable processes. MEMS processes can be very useful but only if the limitations of processes, structures, and materials are known and respected.

The MEMS cycle typically consists of design; mask layout; repeated cycles of deposition, patterning, and etching; interconnection; testing; and design refinement. Development of a new device often takes many of these cycles so the overall development time is more likely to be years rather than months.

True integration of electronics is difficult. Electronic and structural processes are often incompatible so compromises are made in both. Integration often results in low yield. Moreover, changes are time-consuming and expensive.

Packaging is a critical element in any end use. Fascination with microstructures should not overshadow the ultimate goal of producing a useable, packaged device.

Low cost is only achieved with very large volume production. If the market is less than a few hundred thousand per year, the cost per device is likely to range from fifty to several hundred dollars.

All technologies have flaws: Some developers talk about them; many don't. Know the flaws and design accordingly. Conventional thinking hinders development. Merely using MEMS to miniaturize existing structures often results in poor performance.

When to invest in new-MEMS development:

- When long development times are acceptable
- When another critical project is not dependent on a working MEMS device

- When the application has very demanding requirements for:
    - extremely small size or
    - very large numbers of identical devices
- When research in MEMS is specifically desired

When to adapt existing MEMS devices or use conventional fabrication:

- When short development time is required or the budget is limited
- When the device is in the critical path of the project
- When easy availability of components with reliable characteristics is more important than high-volume production
- When exacting performance specifications are more important than very small size
- When demonstration of feasibility provides leverage for future work

Questions to ask when developing sensors:

What is the noise floor in terms of minimum detectable signal? Is the noise floor predicted or measured? What factors were considered if the noise was predicted?

Does the sensor respond to stimuli other than the intended input? Have such spurious responses been considered? If the device is intended to sense acceleration, what is its response to pressure, temperature, or electromagnetic interference?

Will the sensor be evaluated in terms of operation performance goals early in the development cycle?

Can the sensor be packaged to survive the intended environment while maintaining its performance? Has packaging been considered early in the design cycle?

Are laboratory performance measurements representative of the intended operating conditions for the device?

What's the downside? There are always limitations; if none are given, then they've been overlooked.

Lessons Learned: Velocity Sensors

Design of inertial sensors suitable for acoustic velocity sensing is straightforward. Meeting the noise-floor requirement is challenging but there are many candidate sensors suitable for this application. However, producing a sensor that responds accurately to the acoustic particle velocity is a much more difficult problem. The ideal, unconstrained, neutrally buoyant sensor is hard to approach in a practical design.

Essential aspects of velocity sensing:

- The sensor and its mount or suspension must not perturb the acoustic field
- The sensor must respond with good fidelity to the local particle velocity
- The sensor must be isolated from vibration transmitted through the suspension
- The sensor must be protected from non-acoustic fluid motion

The conflicting requirements of permitting the sensor to move freely in the acoustic field but, at the same time, keeping the sensor in the same average location with respect to its mount present the greatest challenge in velocity sensor design.

Required properties of a velocity-sensor suspension:

- The natural frequency of the suspension/sensor system must be well below the signal band of interest
- The suspension stiffness must be isotropic in as many dimensions as the velocity field must be resolved
- Shear stiffness must be considered when using compliant materials
- Periodic properties must be considered when embedding arrays of sensors in compliant material

Velocity sensor testing must account for the suspension and must account for response in directions other than the principal axes of the system. If the actual mount is not used in testing then there is no assurance that the sensor has been characterized adequately. If the only measurements are on-axis response and cross-axis sensitivity, then anisotropy in the directional response will go undetected. Transmissibility from the mount to the sensor must also be evaluated.

Chapter 1

Introduction

Regional actions in littoral waters will likely take the form of rapidly developing conflicts involving multiple surface-ship, mine, and submarine threats. The mission effectiveness of various Navy assets would be greatly enhanced with a deployable multi-threat sensor system for underwater warfare. One critical component of such a system is a miniature, directional hydrophone (or acoustic velocity sensor). While useful in platform systems as well, miniaturization of velocity hydrophones has the greatest payback in deployable systems – either sonobuoy systems or longer-life deployable systems.

The directional sonobuoys used in airborne ASW employ accelerometer-based velocity sensors using piezoelectric materials for transduction, but miniaturization favors other forms of transduction. One form that is particularly suited to microfabrication is the differential-capacitance accelerometer. While this structure is relatively common in medium- to high-grade commercial accelerometers, it has been largely ignored for application to underwater acoustic velocity or acceleration sensors[1].

Beyond sonobuoys, there are a number of emerging applications for underwater acoustic velocity sensors[2]. Some work is being done in the replacement of pressure hydrophones with multi-axis velocity sensors on various platforms and other investigations have focused on two-dimensional arrays of single-axis sensors. While the impetus for this investigation was a substantial upgrade for directional sonobuoys and deployable systems, the velocity-sensor element developed here has application to these other systems.

The differential-capacitance accelerometer is capable of good long-term stability and exceptional low-frequency noise performance. The effective sensor impedance is determined not by the signal frequency but by the AC-drive frequency. The low sensor impedance relaxes the requirements on signal cables. In addition, the critical preamplifier characteristics are those at the AC-drive frequency so low-frequency noise can be avoided. The differential-capacitance accelerometer responds to zero-frequency or DC signals so these devices can be used in self-orienting three-axis sensors for deployable systems. Synchronous detection of the AC-modulated signal yields a high-degree of immunity to external interference and is amenable to straightforward multiplexing and optical or RF transmission.

In contrast to pressure hydrophones, current-technology directional hydrophones occupy a large fraction of the package volume in deployable acoustic systems so there is great value in miniaturization. Small pressure hydrophones are readily available in 4 to 10 cm^3 packages while mass-produced two-axis directional sensors range from 400 to 800 cm^3.

Initially, the project emphasis was on development of an accelerometer-based hydrophone using the electron-tunneling accelerometer pioneered by the Jet Propulsion Laboratory (JPL). A device was demonstrated and an attractive noise floor was achieved in a package of suitable size (Fig. 1.1). However, fabrication proved expensive and labor-intensive

so only a few successful devices were made. Future improvements in MicroElectroMechanical Systems (MEMS) fabrication techniques may permit this device to be produced economically but such development was well outside the scope of this investigation.

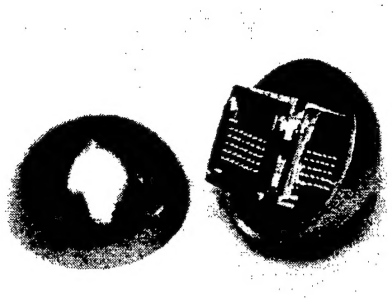

Figure 1.1. Two-axis accelerometer module for use as a directional hydrophone element. The transducers are electron-tunneling devices. The one-inch diameter titanium sphere is the outer pressure housing. The package is neutrally buoyant in water.

Development times for new MEMS devices are often measured in years rather than months. Once it became clear that an alternative to the electron-tunneling accelerometer was needed, the decision was made to search for a commercial product from which the MEMS structure could be removed and adapted for use in the directional hydrophone. While this meant abandoning the self-noise goal, it also enabled a large number of experiments to be run in a relatively short time since there was no development cycle to consider. The first experiments were conducted by buying commercial MEMS-based differential-capacitance accelerometers, carefully cutting them open, and extracting the MEMS chip!

A detailed analysis of the performance of such a chip in a high-frequency AC-drive configuration was done. In particular, a noise budget was developed and, each time a dominant noise source was isolated, that source contribution was reduced until the limiting noise was the intrinsic noise in the chip itself. A crucial conclusion of this analysis was that it is possible to reach the intrinsic chip limit with a practical circuit design.

In FY99, a commercial device manufacturer[3] agreed to sell Penn State University a small quantity of bare MEMS differential-capacitance chips. These structures are normally integrated into commercially marketed accelerometers but the bare structures were well suited to this investigation. The emphasis in miniature commercial accelerometers is not on extremely low noise; consequently, the MEMS chip itself has a substantially lower fundamental self-noise than is achieved in the packaged, commercial product. By using these MEMS chips in a new, full-bridge architecture with specially designed electronics, a substantial reduction in noise floor was achieved. Fundamentally, differential capacitance has a lower responsivity than electron tunneling but the gain in response in electron tunneling is offset by a higher noise floor. More importantly, the fabrication and mass production issues have already been solved for the differential capacitance chip through substantial industry investment. There is no commercial off-the-shelf (COTS) MEMS device that meets the self-noise specification for this project; however, the commercial devices can be improved with better drive and detection electronics and a larger proof mass.

The MEMS chips used here are considerably smaller than the limit imposed by the volume goal for this project. Consequently, once the modifications have proven the potential, the MEMS structures can be scaled *up* in size to produce an accelerometer useable in the directional hydrophone. Increasing the size increases the proof mass, which reduces the limiting molecular-agitation component of the self-noise. For example, the electron-tunneling accelerometer has a proof mass of over 100 milligrams while the differential-capacitance chip has a proof mass of only 2.5 milligrams.

The modifications consisted of one structural modification and substantially different drive and detection circuitry. Instead of using a single chip, two chips were used in a back-to-back arrangement. This partially cancelled the manufacturing offset in the structure and enabled electrical connection as a full bridge.

In the commercial implementation, the differential element is driven with a square wave and the bridge output is amplified and detected with a switching detector. All of the electronics are integrated on a single ASIC (application-specific integrated circuit) in CMOS (complementary metal-oxide semiconductor) and wirebonded to the differential-capacitance chip. CMOS is one of the most economical ASIC processes; however, the noise performance of CMOS amplifiers is relatively poor. Although an exhaustive analysis was not performed on the CMOS chip, the limiting noise is probably either in the first-stage amplifier that is connected to the bridge output or in the switching detector.

Because the permissible budgets for both circuit cost and volume are more generous in the directional hydrophone than in the commercial accelerometer, the ASIC circuitry was completely replaced with surface-mount (small-outline integrated circuit or SOIC) electronics (Fig. 1.2). The result was a two-order-of-magnitude improvement in self-noise. The custom circuitry also enables straightforward multiplexed operation of multiple sensors and self-diagnosis of the health of each sensor element.

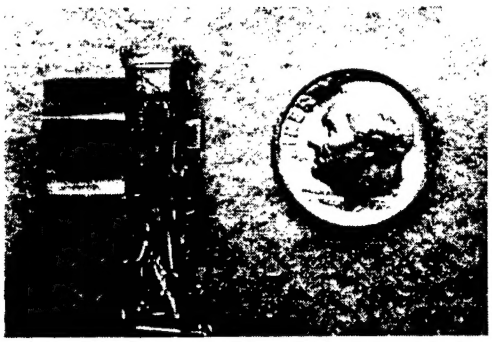

Figure 1.2. Differential-capacitance accelerometer with surface-mount electronics. The two back-to-back MEMS chips are centered inside the black oval. The acrylic stub to the left is a temporary fixture to permit clamping in a calibrator. This device can be mounted in a pressure housing as in Fig. 1.1 or it can be cast directly into syntactic foam for neutral buoyancy and pressure resistance.

For mass production, the differential-capacitance device has several advantages over the electron-tunneling device: (1) it can be designed for structural symmetry to permit balancing of residual stress gradients, (2) clearances and tolerances are more forgiving, (3) no unusually thin structures are required, and (4) the dynamic range is much larger so the capacitive sensor can be operated with unconditionally stable open-loop electronics. In addition, the requirement for high-frequency drive in the differential-capacitance accelerometer leads to straightforward frequency-division multiplexing. In addition, the AC drive can be implemented in such a way as to enable direct integration into a sigma-delta analog-to-digital conversion loop thereby providing direct digital output.

The report on this development is organized as follows. Chapters 2 and 3 discuss fundamental limits on signal detectability in miniature sensors and the associated electronics. These two chapters provide a basis for analysis of sensors in order to find the best-case performance, which is an important first step in sensor design. There is little point in proceeding with a design that cannot meet a specification because of a fundamental limitation.

Chapter 4 treats the limitations in structure introduced by MEMS processing. These aspects do not represent limits in detectability directly; however, structural problems can prevent a promising sensor structure from achieving its potential. Chapters 2 through 4 give an overview of limiting processes that is rather general to miniature sensors of many varieties, although the examples are specific to the accelerometers intended for acoustic velocity sensors.

Chapter 5 reviews the specific development of the directional-hydrophone element that is based on the MEMS differential-capacitance chip. This chapter can be read independently for an overview of the hydrophone work.

Chapter 6 outlines a crucial aspect of acoustic velocity sensing – the sensor suspension. It is insufficient to design a sense element of the required noise-floor, dynamic range, and bandwidth. If the sense element is not suspended properly in the acoustic field, the performance will be poor. The impact of the suspension is vital to understand. Suspension design is as critical to sensor performance as any other aspect concerning the sensor itself. This chapter is relevant to any employment of acoustic velocity sensors and can be read independently.

---

[1] The term "velocity sensor" denotes an inertial sensor that produces an output from which acoustic particle velocity can be determined directly. Geophone-based velocity sensors produce a voltage that is directly proportional to velocity and are true velocity sensors. Many velocity sensors use accelerometers and so measure acoustic particle acceleration but the conversion to velocity is straightforward. Some earlier directional hydrophones measured pressure gradient through differencing of two pressure-sensing elements. These pressure-gradient sensors are not inertial sensors and infer the acoustic particle velocity from the gradient in pressure.

[2] M. Berliner and J. Lindberg, *Acoustic Particle Velocity Sensors: Design, Performance, and Applications*, American Institute of Physics Press, Woodbury, NY, 1996.

[3] CSEM Microsystems, Centre Suisse d'Electronique et de Microtechnique SA, P.O. Box CH-2007, Neuchâtel, Switzerland. Various commercial products are mentioned in this report but none of these references should be construed as endorsements or recommendations.

Chapter 2

Performance Limits: Sensor Self-Noise

Basic Thermal Fluctuations

All mechanical structures are exposed to a continuous spectrum of fluctuating forces associated with molecular vibration. These forces are often unimportant for sensors of conventional size – dimensions of centimeters or more and proof masses of tens of grams or more. At these scales, other sources of noise (notably, preamplifier noise) typically dominate the noise floor. However, as sensor elements are reduced in size, the effects of molecular agitation become increasingly noticeable. At sufficiently small scales, molecular thermal agitation can overwhelm all other sources of noise.

The fluctuating forces result from thermal agitation of molecules. These molecules may be the molecules that comprise the structure itself or they may be the liquid or gas molecules in which the structure is immersed. A body that is otherwise free is subject to a continuous bombardment by surrounding molecules and this can be a crucial design consideration for small bodies. The molecular bombardment forces a free body into an erratic path (see Fig. 2.1) or "random walk."

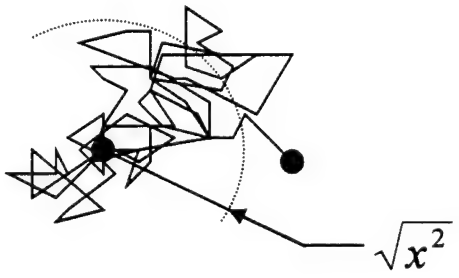

Figure 2.1. Illustration of the random walk of a free particle. The root-mean-square displacement is a measure of the particle's average displacement over some time from the starting point. The actual path is much more irregular than suggested by the rms displacement.

The root-mean-square (rms) displacement in some time, t, is given by the following relation[1]:

$$\sqrt{x^2} = \sqrt{2Dt} \tag{2.1}$$

where

$$D = k_B T / R_{mech} \tag{2.2}$$

and k_B is Boltzmann's constant (1.38×10^{-23} joules per kelvin), T is the absolute temperature, and R_{mech} is the mechanical resistance. The mechanical resistance is the ratio of force applied to velocity produced. For a sphere of radius, a, in a viscous fluid[2] with viscosity, η,

$$R_{mech} = 6\pi\eta a \tag{2.3}$$

The rms displacements for one second and various sized spheres (having the density of water) immersed in water are given in Table 2.1.

Radius, a [m]	Drift, x [m]	Drift [body lengths]
1	$7 \cdot 10^{-10}$	$7 \cdot 10^{-10}$
10^{-3}	$2 \cdot 10^{-8}$	$2 \cdot 10^{-5}$
10^{-6}	$7 \cdot 10^{-7}$	0.7
10^{-9}	$2 \cdot 10^{-5}$	$2 \cdot 10^{4}$

Table 2.1. Relationship between the size of a body and the one-second random-walk drift at room temperature in water. The drift is given in both meters and in body lengths.

Even a body of millimeter size is influenced sufficiently that the motion needs to be considered in designing a high-sensitivity sensor. A body of micrometer size is disturbed by nearly its own dimension per second and a body of nanometer size is thrown about on scales much larger than its own dimension.

The free-body problem is not directly applicable to most sensor structures, however. A sensor may have a diaphragm or a proof mass that is intended to move but within some support constraint. A useful first model for such a structure is the damped, mass-spring oscillator. The moving part of such a structure is still influenced by molecular bombardment and by internal molecular agitation but the time-average position does not change. A schematic diagram of such an oscillator with its response to an impulsive input is shown in Fig. 2.2.

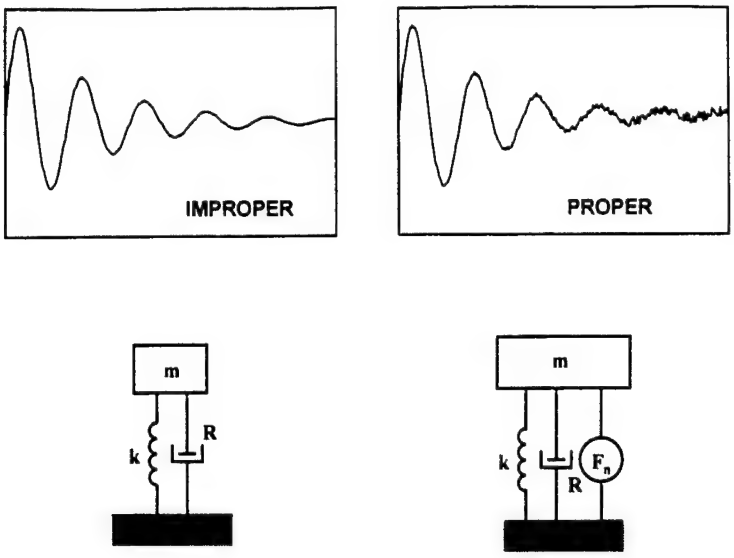

Figure 2.2. Damped mass-spring oscillator with an impulsive input. The system on the left is unrealizable; its amplitude decays to arbitrarily small values with sufficient time. The system on the right decays into the background of molecular-vibration induced fluctuations.

The differential equation of motion of the damped oscillator is often written as

$$m\ddot{x} + R\dot{x} + kx = 0 \tag{2.4}$$

where m is the mass, R is the mechanical resistance, k is the spring constant, and x is the displacement from equilibrium. This equation is, in fact, wrong. It predicts a violation of the Second Law of Thermodynamics. According to this equation, the amplitude of the oscillation will decay to any arbitrarily small value regardless of the temperature of the system. Because the system is made of thermally agitated molecules and the system is in contact with other thermally agitated molecules, its amplitude can only be made arbitrarily small by reducing the temperature. The homogeneous differential equation suggests that the oscillator can self cool below ambient!

The proper equation of motion for a damped, mass-spring oscillator is

$$m\ddot{x} + R\dot{x} + kx = f_n(R,T) \tag{2.5}$$

Here, the right-hand side is not zero. There is a forcing function, f_n, with a value dependent on the mechanical resistance and temperature, T. This forcing function represents the thermal agitation of the system. After an initial disturbance damps out, the forcing function maintains the system at a level of fluctuation consistent with thermal equilibrium.

It is significant that the forcing function is directly associated with the mechanical damping. If there is a path by which energy can leave a system, then energy can return to the system from the environment along the same path. Energy loss from the system's point of view

is sometimes called loss or damping and, in other instances, is called radiation (see Fig. 2.3). *Any mechanism of connection between the system and environment provides a route for the equilibrium thermal fluctuations in the environment to act on the system.* In equilibrium, the rate of energy leaving the system through the "loss" mechanisms is balanced, on average, by the rate of energy entering the system from the environment. This intimate association between fluctuations and loss mechanisms forms the basis for the Fluctuation-Dissipation Theorem[3].

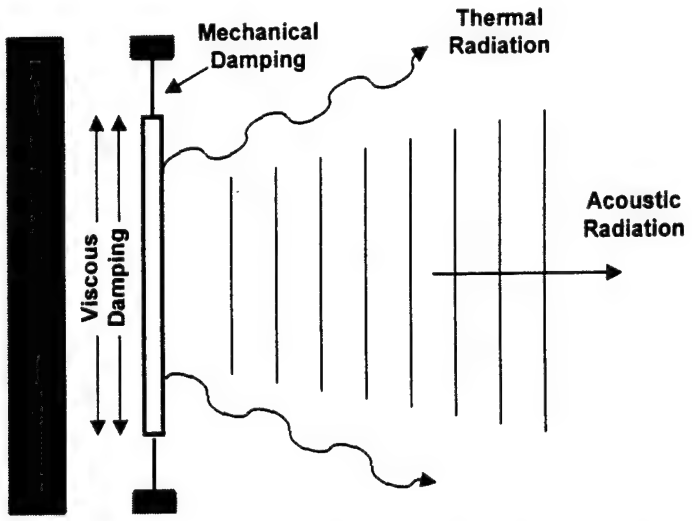

Figure 2.3. Illustration of various mechanisms for energy "loss" from a moving diaphragm (a condenser microphone, for example). Any mechanism that provides a path for energy from the structure to the environment also provides a path through which fluctuations in the environment can influence the structure.

While the essence of the theorem is contained in the description given above, there are a number of ways of expressing the theorem quantitatively. One way is to relate the autocorrelation of the fluctuating force, F, to the mechanical resistance as follows[4]:

$$R_{mech} = \frac{1}{2k_B T} \int_{-\infty}^{\infty} \langle F(0)\, F(\tau) \rangle \, d\tau \tag{2.6}$$

Here the angle brackets indicate an average over many sets of interactions. However, there is a more useful form derived by Nyquist[5]. The spectral density of the fluctuations can be related directly to the damping. In terms of mechanical resistance, the mean-square force is

$$F_n^2 = 4k_B T\, R_{mech}\, df \tag{2.7}$$

and, in terms of electrical resistance, the mean-square voltage, is

$$V_n^2 = 4k_B T\, R_{elect}\, df \tag{2.8}$$

The latter expression is the familiar Johnson noise expression[6]. In these expressions, V_n^2 and F_n^2, are mean-square values. These will often be written as spectral densities, in which case the left-hand-side quantity would be force-squared per hertz and the *df* factor (the increment of bandwidth) would not appear on the right-hand side. Another alternate form for either relation is obtained by dividing by R^2. The alternate form relates fluctuations in either velocity or current to the resistance.

An important feature of the Nyquist relations is that they predict the frequency distribution of the fluctuations. If R is a function of frequency, then the spectral density has the same functional dependence on frequency. If R is independent of frequency, then the fluctuations are white up to a very high frequency set by quantum-mechanical limits. There are important cases in which R is a function of frequency: for example, mechanical or electrical loss tangent or radiation resistance. The Nyquist relations are still valid.

The limiting behavior at very high frequency (the quantum-mechanical limit) is of little consequence for sensors operating above a few kelvin in temperature. The complete expression of the Nyquist relation is

$$F_n^2 = 4k_BT\left[\frac{hf/k_BT}{e^{hf/k_BT}-1}\right]R_{mech}\,df \tag{2.9}$$

where h is Plank's constant (6.63×10^{-34} joule-sec). Even at 1 K and 1 GHz, the factor in square brackets is 0.98. At higher temperatures and lower frequencies, it can be omitted entirely.

Ultimately, it is not the absolute level of noise that is critical. Instead, the critical measure is the achievable signal-to-noise ratio. A very low noise level is of no value if the response to signals is correspondingly weak. Conversely, a high noise level is not necessarily bad if the sensor compensates with a large signal response. One technique for avoiding a skewed emphasis on either signal or noise level is to describe the noise in terms of the noise-equivalent signal (NES). By describing the noise in terms of the signal level that would produce the same level as the noise, the performance of the sensor is directly tied to the minimum detectable signal. Specifying the noise of an accelerometer, for example, as 25 nanovolts per root hertz is incomplete. If the acceleration response is 1 volt per g, then the noise-equivalent-acceleration[7] (NEA) is 25 nano-g per root hertz. The NEA can be compared directly to the expected signals of interest or to ambient noise levels.

For a simple accelerometer (one having a single mechanical degree of freedom), the NEA with respect to molecular-thermal noise is found by setting the mean-square value of the inertial force, ma_n, associated with the NEA, a_n, equal to the molecular-thermal fluctuation in mean-square force given by the appropriate Nyquist relation:

$$(ma_n)^2 = 4k_BT\,R\,df \tag{2.10}$$

Solving for the spectral density of the acceleration noise and using the definition of the simple-harmonic oscillator's quality factor, $Q = \omega_0 m/R$, the noise-equivalent acceleration is[8]

$$\frac{a_n^2}{df} = 4k_B T \frac{R}{m^2} = 4k_B T \frac{\omega_0}{mQ} \tag{2.11}$$

If the mechanical resistance, R, is constant, then the spectral density drops as the inverse of the square of the proof mass. If the resonance frequency ($f_0 = 2\pi\omega_0$) and the Q are kept constant, then the spectral density drops as the inverse of the proof mass. In either case, miniaturization raises the noise floor: as the sensor size is decreased, the mass decreases. Furthermore, if the mechanical resistance is primarily the result of viscous losses in the gas surrounding the moving elements (as it often is in thin-film microstructures), then the mechanical resistance may increase with size reduction, making the NEA grow even faster.

This is the first and most fundamental limit associated with miniaturization. Even high-quality microphones are limited by molecular-thermal agitation over parts of their operating bands[9] so the size scale does not have to be excessively small for these effects to be important.

Eq. 2.11 is a fundamental design equation for miniature accelerometers. The NEA can be decreased by decreasing the resonance frequency (which affects the sensor bandwidth), by increasing the proof mass (necessitating an increase in stiffness to maintain the same resonance), or by increasing the Q. In an accelerometer (in contrast to a pressure sensor), the structure can be operating in vacuum and the Q can, in principle, be very high. This may lead to problems in dynamic range and impulse response, however, so there are practical limits to this strategy.

If the proposed sensor does not have a sufficiently low NEA according to this consideration of thermal fluctuations, then it must be redesigned. Of course, in practice it may be very difficult to reach the fundamental limiting NEA, so other sources of noise must normally be considered, too.

Noise from Molecular Collisions

For normal internal gas pressures and conventional size scales, the noise from molecular collisions is modeled well by Nyquist's relation and the system's mechanical resistance. It is instructive, however, to consider the collision phenomenon from a different perspective, which will shed some light on the problem at micrometer scales. First, a brief summary of shot noise is given and then, the connection to molecular collisions is developed. There is an important transition region that is reached on the micrometer scale. Furthermore, this development provides a foundation for the electronic-noise treatment in the next chapter.

Shot noise is noise associated with flow of discrete "carriers" for the case in which those carriers act independently. As an example, consider the flow of electrons. On a sufficiently fine time scale, the flow appears to be an irregular stream of particles, each with a fixed charge, q (see Fig. 2.4).

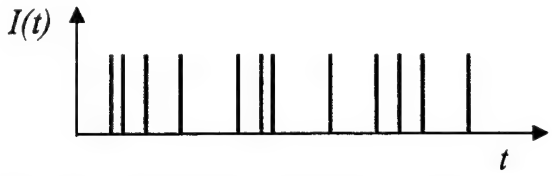

Figure 2.4. Sequence of impulses representing current flow as individual charges arriving at random times.

In terms of this stream of particles, the current, $I(t)$, can be written as the sum of delta functions:

$$I(t) = q\sum_{i=1}^{\infty}\delta(t-t_i) \tag{2.12}$$

This expression for the current can be expanded as a Fourier series,

$$I(t) = \sum_{k=0}^{\infty}\left[a_k\cos(2\pi f_k t) + b_k\sin(2\pi f_k t)\right] \tag{2.13}$$

where the a_k coefficients are as follows:

$$a_k = \frac{2}{\mathsf{T}}\int_0^{\mathsf{T}} I(t)\cos(2\pi f_k t)dt = \frac{2q}{\mathsf{T}}\sum_{i=1}^{N}\cos(2\pi f_k t_i) \tag{2.14}$$

and the b_k coefficients are given by similar expressions with the cosines replaced by sines. The expansion is taken over a period, T, that encompasses many (N) events. Consequently, each Fourier component covers a frequency interval, $\Delta f = 1/\mathsf{T}$. The mean-square value of the current in one of those intervals is

$$\overline{i_k^2} = a_k^2\,\overline{\cos^2(2\pi f_k t)} + b_k^2\,\overline{\sin^2(2\pi f_k t)} + a_k b_k\,\overline{\cos(\,)\sin(\,)} = \frac{1}{2}\left(a_k^2 + b_k^2\right) \tag{2.15}$$

where the overbars indicate time averages. The time average of either cosine-squared or sine-squared is ½ whereas the time average of the sine-cosine product is zero. Using Eq. 2.14 for a_k and the similar form for b_k, this can be rewritten as two sums, one involving the products of factors with identical indexes and the other involving the products of factors with different indexes:

$$\frac{1}{2}\left(a_k^2 + b_k^2\right) = \frac{2q^2}{\mathsf{T}^2}\left\{\sum_{i=j}(\,) + \sum_{i\neq j}(\,)\right\} \tag{2.16}$$

where

$$\sum_{i=j}(\,) \;=\; \sum_{i=1}^{N}\left[\cos^2(2\pi f_k t_i) \;+\; \sin^2(2\pi f_k t_i)\right] \;=\; N \tag{2.17}$$

and

$$\sum_{i\neq j}(\,) \;=\; \sum_{i\neq j}^{N}\left[\cos(2\pi f_k t_i)\cos(2\pi f_k t_j) \;+\; \sin(2\pi f_k t_i)\sin(2\pi f_k t_j)\right] \tag{2.18}$$

The second summation is zero *if and only if the carriers are independent* – that is, if the arrival of one is not in any way dependent on the arrival of any other. In that case,

$$\frac{1}{2}\left(a_k^2 + b_k^2\right) \;=\; \frac{2q^2 N}{\mathsf{T}^2} \tag{2.19}$$

Since the average current and the frequency interval are given by

$$\bar{I} \;=\; \frac{qN}{\mathsf{T}} \qquad ; \qquad \Delta f \;=\; \frac{1}{\mathsf{T}} \tag{2.20}$$

the current fluctuation has the classic shot noise form:

$$\overline{i_k^2} \;=\; 2q\bar{I}\Delta f \tag{2.21}$$

To reiterate, this expression is valid for processes consisting of events that are impulse-like and independent. If the events are not impulse-like, the impulse response can be introduced into Eq. 2.12 in place of the delta functions. If the events are not independent, then the summation, Eq. 2.18, involving the cross terms must be computed. It is critically important to remember that the shot-noise expression only applies to processes involving independent events. Photons incident on a photodetector are independent; electrons flowing through a metallic conductor are not.

There is a more general way to write shot noise expressions:

$$s_n^2 \;=\; 2[c]\frac{S_0}{A}\Delta f \tag{2.22}$$

This expression relates the mean-square fluctuation, s^2, in some flux density to the quantity, c, per carrier, the average flux density, S_0, the area, A, through which the flux is passing, and the frequency interval.

For electrical current, the quantity per carrier is the charge, q, and the flux density is the current flux density, J:

$$j_n^2 = 2[q]\frac{J_0}{A}\Delta f \tag{2.23}$$

If this equation is multiplied by A^2, then Eq. 2.21 results. For photon flow, the quantity per carrier is the energy, hf, and the flux density is the optical intensity, $\mathbf{I}$:

$$\mathbf{I}_n^2 = 2[hf]\frac{\mathbf{I}_0}{A}\Delta f \tag{2.24}$$

For molecular collisions, the quantity per carrier is the average momentum transferred per collision, $2mv$, and the flux density is the momentum flux density, which is the pressure:

$$p_n^2 = 2[2mv]\frac{P_0}{A}\Delta f \tag{2.25}$$

This last form – the one for pressure fluctuation – warrants further development. It contends that the mean-square pressure fluctuations are directly proportional to the static pressure, P_0, a contention that is not consistent with measurement!

Consider the collision and rebound of a molecule from a surface as in Fig. 2.5. The average speed of the molecule is v and the surface is moving with some speed, $\dot{x}$. (The collision is shown as an elastic, specular reflection but this is a simplification. In actuality, a molecule that collides with the surface under normal conditions, rebounds with random direction. Because the distributions of speed and direction in the gas do not change with time, there will always another molecule that does rebound in the direction specular to the incident molecule. For the purposes of simplifying the analysis, we will pair up the appropriate incident molecule with the appropriate rebounding molecule as if the collision occurred as shown in the figure.)

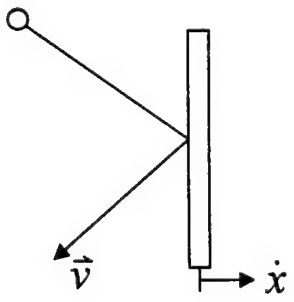

Figure 2.5. Collision of a single molecule with a moving surface. The forces that result depend on the velocity distribution of the molecules and whether or not there is significant interaction between incoming and rebounding molecules.

The net force on the surface is the sum (over all collisions) of the rate of change in momentum of the molecules that strike the surface. This is equal to the molecular flux toward the surface times the momentum change per collision or,

$$F = \frac{n}{2} A(v_x - \dot{x}) \cdot 2m(v_x - \dot{x}) \tag{2.26}$$

Here, the relative speed between the molecule and the surface is given by the difference between the x-component of the molecule's speed and the speed of the surface. The molecular flux is the number density, n, of molecules times the area, A, times the relative speed. The factor of ½ is used because only half of the molecules are initially traveling toward the surface. The momentum change per collision is two times the molecular mass, m, times the relative speed since the molecule changes direction in the collision.

Expanding the product produces three terms:

$$F = nm\overline{v_x^2} A - 2nm\overline{v_x} A\dot{x} + nm\dot{x}^2 A \tag{2.27}$$

If the speed of the surface is much less than the average molecular speed, the third term can be ignored. The first term is the kinetic-theory expression for static pressure times area and the second term is some factor times the speed of the surface; therefore, the second term is equivalent to a resistance:

$$F = P_0 A - R_{mech}\dot{x} \tag{2.28}$$

where

$$R_{mech} = 2nm\overline{v_x} A \tag{2.29}$$

Using the mechanical resistance, the Nyquist relation gives the fluctuations in the force on the surface:

$$F_n^2 = 4k_B T R_{mech}\, df = 8nmk_B T\overline{v_x} A\, df \tag{2.30}$$

Kinetic theory also gives the following relations[10]

$$P_0 = nk_B T \quad ; \quad \bar{v} = 2\overline{v_x} \tag{2.31}$$

which, when substituted into Eq. 2.30, give the same result obtained from the generalized shot-noise form (Eq. 2.25):

$$p_n^2 = \frac{F_n^2}{A^2} = 2[2m\bar{v}]\frac{P_0}{A} df \tag{2.32}$$

We should be able to calculate the pressure fluctuations another way, also. If the surface moves very slowly compared to the molecular speed, we can use Stokes' expression[11] for resistance. The mechanical resistance associated with oscillatory motion of a circular disk in a viscous fluid is given by:

$$R_{mech} = 16\eta a \tag{2.33}$$

where η is the viscosity. Using Nyquist's relation, the corresponding pressure fluctuation would be

$$p_n^2 = 4k_B T 16\eta a df / A^2 \tag{2.34}$$

Eq. 2.32 indicates that the pressure fluctuations should be directly proportional to static pressure, whereas, Eq. 2.34 shows virtually no dependence on static pressure (viscosity is a very weak function of pressure at normal pressures). Which expression is correct?

For the shot-noise form to be correct, the molecular collisions must be independent. *At normal pressures, gas-molecule collisions are strongly dependent.* The average distance between collisions (the mean-free-path) is very small (about 0.1 micrometer at one atmosphere) so a molecule from the bulk of the gas cannot reach the surface without interacting with many molecules that are rebounding from the surface. The rebounding molecules carry information regarding the surface motion and, in effect, transfer that information to the incoming molecules. The motion of the surface modifies the velocity distribution through these interdependent collisions.

Consequently, the shot-noise expression is not correct for normal pressures. While the derivation of Eq. 2.32 did not explicitly introduce shot noise, no accounting was made for the change in velocity distribution, which is equivalent to assuming independence of collisions. At normal pressures, Eq. 2.34 is the correct expression and measurements corroborate this expression.

The viscosity can be expressed in terms of kinetic theory[12]:

$$\eta = nm\bar{v}l_{mfp}/3 \tag{2.35}$$

where l_{mfp} is the molecular mean-free-path. The ratio of the mean-square pressure based on shot theory (Eq. 2.32, doubled to account for collisions on both sides) to the mean-square pressure based on viscous resistance (Eq. 2.34) is

$$\frac{p_{n1}^2}{p_{n2}^2} = \frac{3\pi}{8}\frac{a}{l_{mfp}} \tag{2.36}$$

The two expressions are approximately equal when the mean-free-path equals the disk radius. If the mean-free-path is sufficiently large compared to the dimensions of the structure, then the

collisions do act independently and the shot analysis is correct. This is the regime of free-molecular (or Knudsen) flow. Overall, the pressure fluctuations are dependent on static pressure in the manner shown in Fig. 2.6.

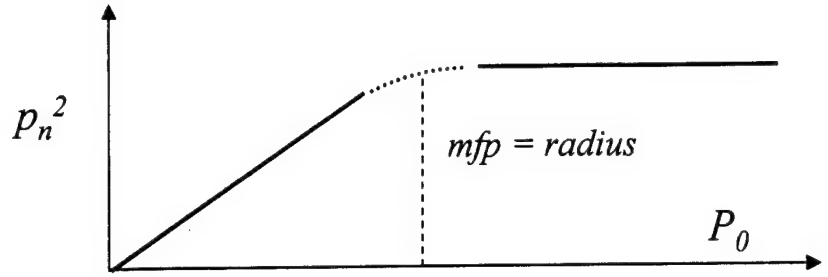

Figure 2.6. Transition in pressure fluctuations from molecular collisions. Mean-square fluctuations are proportional to static pressure in the free-molecular regime to the left but are independent of static pressure in the normal viscous-flow regime to the right. The transition takes place when the molecular mean-free-path (mfp) is about the same length as the significant dimension of the surface.

For low pressures (mean-free-path large compared to the structure), the fluctuations are directly proportional to the static pressure. For pressures under which a normal, viscous-flow model is appropriate, the mean-square fluctuation is only dependent on pressure to the degree that the viscosity is pressure dependent.

At one atmosphere, the molecular mean-free-path is about 0.1 micrometer but the effects of collision independence extend to somewhat higher pressures because of a more subtle effect. At low pressure (i.e., large mean-free-path) but before fully developed free-molecular flow, the assumption of zero-slip at boundaries is not accurate[13]. This "slip-flow" region causes the transition to free-molecular behavior to start at smaller mean-free-paths than expected. For example, there is about a ten-percent effect at dimensions of 6 micrometers for a 0.1-micrometer mean-free-path. While the analysis above was done for an isolated plate, similar results are obtained for flow between surfaces. For surface-micromachined structures, gaps of 2 micrometers or less are common. Consequently, some degree of free-molecular flow effects can be expected even at one atmosphere in such structures. In addition, the damping is so large at one atmosphere in these structures that they are often designed to be operated with some vacuum. This would necessitate treatment as free-molecular flow. In any event, it is necessary to understand the distinction between dependent and independent collisions if shot-theory is to be applied.

---

[1] A. Einstein, *Investigations on the Theory of the Brownian Movement*," Dover, NY, 1956.

[2] L. Landau and E. Lifshitz, Fluid Mechanics, Pergamon, NY, 1959, §20.

[3] H. Callen and T. Welton, "Irreversibility and generalized noise," Phys. Rev. **83**, 34-40, 1951.

[4] F. Reif, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, NY, 1965.

[5] H. Nyquist, "Thermal agitation of electric charge in conductors," Phys. Rev. **32**, 110-113, 1928.

[6] J. Johnson, "Thermal agitation of electricity in conductors," Phys. Rev. **32**, 97-109, 1928.

[7] The "g" is not a primary SI unit but it is far more commonly used with respect to accelerometers than the m/s^2. As a result, the "g" will be used here but the normal unit prefixes will be spelled out to differentiate from the gram. For example, nano-g is acceleration; ng is nanogram.

[8] M. Usher, I. Buckner, and R. Burch, "A miniature wideband horizontal-component feedback seismometer," J. Physics E: Sci. Instruments **10**, 1253-1260, 1977.

[9] V. Tarnow, "The lower limit of detectable sound presure," J. Acoust. Soc. Am. **82**, 379-381, 1987.

[10] F. Sears and G. Salinger, *Thermodynamics, Kinetic Theory, and Statistical Mechanics*, Addison-Wesley, NY, 1975.

[11] H. Lamb, *Hydrodynamics*, Dover, NY, 1945, §339.

[12] P. Morse, *Thermal Physics*, W. A. Benjamin, NY, 1969.

[13] E. Kennard, *Kinetic Theory of Gases*, McGraw-Hill, NY, 1938, Chapter 8.

Chapter 3

Performance Limits: Sensor-Amplifier Interaction

In the last chapter, some of the consequences of sensor miniaturization were considered with regard to the intrinsic self-noise of the sensor. It is also important to consider the limits imposed by the electrical interface. Miniaturization often produces devices of very small capacitance or otherwise acute sensitivity to electrical interconnection length. Often, it is suggested that these problems be solved by integrating the electronics onto the same chip as the mechanical structure. The solution is rarely simple, though. The electronics processes that are compatible with mechanical processes are often not conducive to production of low-noise electronics and the yield of a composite electrical/mechanical chip is often substantially lower than the overall yield in separate production. Therefore, it is reasonable to consider the limits imposed by the electronics. As before, we will start with fundamental limits and work toward more specific cases.

Fundamentals

Consider first, the flow of current through a potential barrier (such as a semiconductor PN junction) as shown in Fig. 3.1.

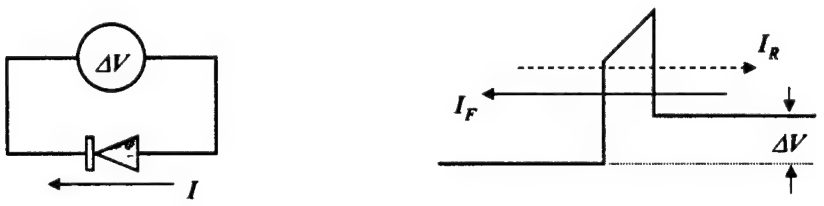

Figure 3.1. Current flow through a potential barrier. On the left, a voltage is applied to a semiconductor diode. The schematic representation of the potential barrier associated with the semiconductor junction and the forward and reverse currents is shown on the right.

The flow can be decomposed into two counter flows, a forward current, I_F, from higher potential to lower ("downhill") and a reverse current, I_R, from lower potential to higher. These two flows are related by the standard Boltzmann function from statistical mechanics[1]:

$$I_R = I_F \, e^{-\frac{q\Delta V}{k_B T}} \tag{3.1}$$

where ΔV is the voltage applied to the junction and q is the charge on an electron. The quantity, $q\Delta V$ is the potential energy associated with the change in electrical potential from one side of the junction to the other and $k_B T$ is the average thermal energy.

The observable current, I, is the difference between the forward and the reverse currents:

$$I = I_F - I_R = I_R \left[e^{\frac{q\Delta V}{k_B T}} - 1 \right] = \frac{\Delta V}{R} \tag{3.2}$$

which is also written as the voltage drop divided by the effective junction resistance, R.

The charge carriers associated with current flow across a potential barrier are independent. Therefore, both the forward and the reverse currents contribute a shot-noise component, which add in the mean-square sense:

$$i_n^2 = 2q I_F \, df + 2q I_R \, df = 2q I_R \left[e^{\frac{q\Delta V}{k_B T}} + 1 \right] df \tag{3.3}$$

Eliminating I_R between Eqs. 3.2 and 3.3 yields the following expression for the current fluctuation:

$$i_n^2 = 2\frac{q\Delta V}{R} \left[\frac{e^{\frac{q\Delta V}{k_B T}} + 1}{e^{\frac{q\Delta V}{k_B T}} - 1} \right] df \tag{3.4}$$

In the limit of small applied voltage,

$$i_n^2 \rightarrow 4k_B T \frac{df}{R} \qquad \text{for} \quad \Delta V << k_B T/q \tag{3.5}$$

The fluctuations are given by the Nyquist expression for current (Eq. 2.8 divided by R^2). In the limit of large applied voltage,

$$i_n^2 \rightarrow 2q I \, df \qquad \text{for} \quad \Delta V >> k_B T/q \tag{3.6}$$

which is the classic shot-noise equation. (At room temperature, the quantity $k_B T/q$ is about 25 millivolts.) As in the case of molecular collisions, we see a transition between a Nyquist expression and a shot expression.

For systems in thermodynamic equilibrium, the Nyquist expression must appear regardless of our approach to the calculation. Above, we started with shot-noise expressions but in the equilibrium limit (small current flow), the Nyquist expression was obtained. In many cases, systems that are not in thermodynamic equilibrium have the Nyquist (or Johnson) noise component but they also have additional noise. Such is the case here where the Nyquist form gives the lower limit.

Transistor Noise

With the framework developed above, consider the fundamental noise associated with a transistor. The bipolar junction transistor (BJT) will be considered first with the results extended to the field-effect transistor (FET). The schematic symbol, the representative structure, and one of many performance models for the BJT are shown in Fig. 3.2.

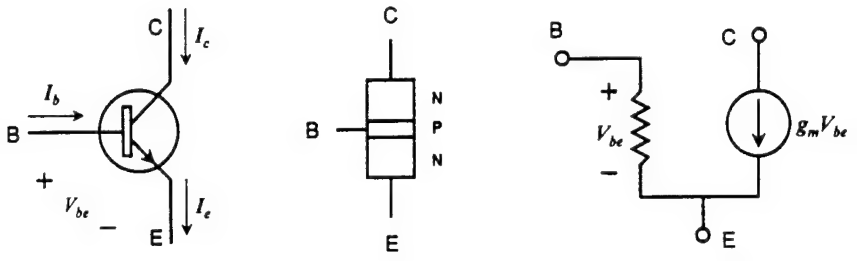

Figure 3.2. Structure of the bipolar junction transistor. The schematic symbol is shown on the left. The physical structure of an NPN transistor is shown in the middle – a thin P-type region separates two N-type regions forming two junctions. One performance model is shown on the right. In this model, the collector current is directly proportional to the base-emitter voltage.

This transistor behaves as a valve in which the bulk of the current flows from the collector (C) to the emitter (E) while being controlled by the voltage between the base (B) and the emitter. The model on the right indicates that the current from collector to emitter is directly proportional to the voltage, V_{be}. The proportionality constant is the transconductance, g_m.

In normal operation, the collector current is very nearly equal to the emitter current and is also much larger than the base current by a factor, β, the current gain:

$$I_E \approx I_C = \beta I_B \qquad (\beta >> 1) \tag{3.7}$$

Typically, β is greater than 100.

Since the current flow crosses a potential barrier in the transistor structure, the collector current can be written as in Eq. 3.2 with I_R being the (unobservable) reverse current through the base-emitter junction:

$$I_C = I_R \left[e^{\frac{qV_{BE}}{k_B T}} - 1 \right] \tag{3.8}$$

In transistor analysis, this equation is known as the Ebers-Moll Equation[2]. The transconductance is the rate of change of collector current with respect to base-emitter voltage.

Differentiating Eq. 3.8 and assuming that the base-emitter voltage is much greater than 25 millivolts yields

$$\frac{dI_C}{dV_{BE}} = \frac{qI_C}{k_B T} \equiv \frac{1}{r_e} \equiv g_m \tag{3.9}$$

(The effective emitter resistance, r_e, is often used instead of the transconductance.)

Since we are interested in comparing the noise introduced by the electronics to the self noise of the sensor, we will consider the noise referred to the input side (the base) of the transistor. From this point on, the *df* in the mean-square noise expressions will be dropped and the quantities will be taken to be spectral densities. There are two components, one component of shot noise related to the base current:

$$i_n^2 = 2qI_B = 2qI_C/\beta \tag{3.10}$$

and one component of Nyquist (Johnson) noise associated with the effective emitter resistance:

$$e_n^2 = 4k_B T r_e = \frac{(2k_B T)^2}{qI_C} \tag{3.11}$$

These quantities are called the equivalent input current noise (spectral density) and the equivalent input voltage noise of the transistor, respectively. Their units are amperes-squared per hertz and volts-squared per hertz. More commonly, the square root is taken. Then the units become amps per root hertz and volts per root hertz. The use of the square root of hertz in the unit causes some confusion but it can be viewed simply as an artifact of taking the square root of the spectral density.

Notice that r_e is not associated with a particular piece of material with some resistivity but it is still a legitimate loss mechanism: the associated current and voltage are in phase. In evaluating Johnson noise, we don't need physical details. It is sufficient to take the real part of the electrical impedance. Notice also that, if we know the current gain of the transistor and the operating collector current, we can predict both components of the noise.

A field-effect transistor (FET) has a transconductance that is the reciprocal of an effective "gate" resistance. This gate resistance produces a voltage noise. Moreover, there is a gate leakage current which produces a shot-like current noise. The g_m and gate leakage are not as easily expressed in fundamental quantities for the FET but they are given in the manufacturers' data sheets. FETs typically have lower g_m and much lower gate current than BJTs.

The equivalent input voltage and current noise components are sufficient to describe the noise behavior of the transistor (and, as discussed below, these same components are sufficient to describe the noise of preamplifiers in general). However, these two quantities can be combined to form other quantities that offer additional insight into the noise performance. One such

quantity is the the product of the voltage and current noise normalized by the Nyquist factor, $4k_BT$:

$$\frac{e_n i_n}{4 k_B T} = \frac{1}{\sqrt{2\beta}} \tag{3.12}$$

This "normalized noise power" is a measure of the quality of the device from the standpoint of noise. The smaller the normalized noise power, the better. Since β is generally less than 1000, this factor would rarely be less than 0.02 for a bipolar device. Because the leakage current is generally much lower in a field-effect transistor, the equivalent current noise for a FET is much lower even though the voltage noise is usually a bit higher. Overall, the factor in Eq. 3.12 may be substantially lower for a field-effect transistor (especially a junction field-effect transistor or JFET), in some cases below 0.001. The significance of this factor will be explored in more detail below.

Another number of some utility is the ratio of the noise voltage to noise current, sometimes called the effective noise resistance:

$$\frac{e_n}{i_n} = \sqrt{2\beta}\left(\frac{k_B T}{q I_C}\right) \tag{3.13}$$

As we will see below, this is the value of a purely resistive sensor for which the ratio of the amplifier's noise contribution to the intrinsic Johnson noise contribution from the sensor would be smallest. Sometimes amplifiers are selected solely on the degree to which the effective noise resistance matches the impedance of the sensor. This can, however, lead to poor amplifier choice. The smaller the noise power (Eq. 3.12), the wider the range of source resistance over which the amplifier's contribution is negligible. A transistor (or amplifier) with a large noise power but accurately matched in terms of noise resistance may be a poor choice compared to another device with much smaller noise power but not well matched for noise resistance. *A single figure-of-merit is inadequate to express the worth of an amplifier in combination with a sensor.* Furthermore, many sensors are not resistive. Capacitive sensors, for example, have an impedance that varies with frequency.

Before examining more specific sensor/amplifer interaction problems, the equations for e_n and i_n (Eqs. 3.10 and 3.11) lead to an important fundamental limitation to electronics that is often overlooked in the design of miniature sensors (especially microfabricated sensors with integrated electronics). The normalized noise power (Eq. 3.12) is roughly constant for a particular device. The only dependence is on the current gain, which is weakly dependent on operating current. (For example, the current gain might change by 50% for a three order-of-magnitude change in the collector current.) The collector current can be increased to reduce e_n but then i_n will increase (or vice versa). Even from device to device, low current-noise devices tend to have high voltage noise; low-voltage noise devices tend to have high current noise. Notice also that the voltage noise is dependent only on collector current and fundamental quantities (Eq. 3.11).

Furthermore, consider the design of a low voltage-noise amplifier. A good, low voltage-noise amplifier might have an input voltage noise of one nanovolt per root hertz. The amplifier (as an operational amplifier) would typically have two input transistors operating differentially with noise voltages that add in the mean-square sense. Therefore, each transistor would need to operate at 0.7 nanovolt per root hertz. From Eq. 3.11, the collector current for each would be 0.8 mA and the total current would be 1.6 mA for the input transistor pair.

What if the desired input voltage noise were 0.1 nanovolts per root hertz? Now the total collector current just for the input transistor pair would be 0.16 amps! In addition, the equivalent emitter resistance for each transistor would be 0.3 ohms. There are two difficult design obstacles here. One is the problem of heat dissipation associated with the steady current of 0.16 amps; the other is the problem of designing the semiconductor structure so that its internal resistance is below the 0.3 ohm equivalent emitter resistance (otherwise, the Johnson noise would increase). Both of these problems can be mitigated only by increasing the physical size of the transistors, resulting in, among other things, very large input capacitance.

This issue is often lost in the enthusiasm for microcircuitry. Modern microprocessors have thousands of transistors per square millimeter and it's usually assumed that any electronics in a MEMS device would occupy a negligible amount of area on the chip. Transistors intended for digital circuitry, however, only have to distinguish reliably between two states – the dynamic range of any one transistor can be remarkably small. This is in stark contrast to low-noise analog circuitry. The area required for a low-noise analog interface amplifier could be significantly *larger* than the mechanical structure. Consider the Analog Devices AD743, a low-noise operational amplifier. Its die outline is shown in Fig. 3.3. The input transistor pair occupies more than one square millimeter!

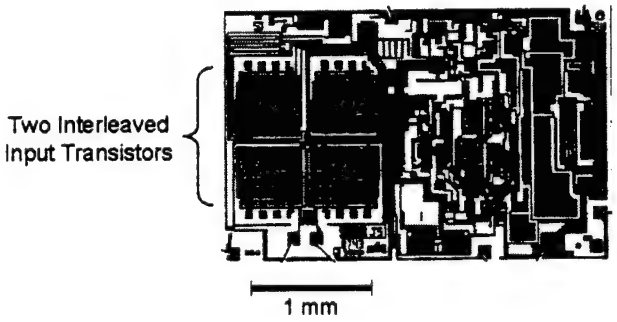

Figure 3.3. Die layout for the AD743 low-noise operational amplifier. The pattern of four squares on the left side of the die is the area required for the two, interleaved input transistors.

Fig. 3.3 illustrates another important aspect of high-performance design. The input transistor pair actually appears to be four elements. In fact, the two transistors are interleaved; geometrically they have a common centroid. Common-centroid design confers two advantages in a practical device. First, the effects of temperature gradients are minimized – a linear gradient affects each transistor on average in the same way. Second, the effects of nonuniformities in fabrication processing are minimized. The processes of etching, ion implantation, and the like are not uniform over millimeter scales. The impact of processing nonuniformity is minimized by the common-centroid design. In contrast, microsensor mechanical structures are not ordinarily

suited to common-centroid design and are sensitive to process irregularities and thermal gradients.

Integration of electronics of this quality onto a micromachined sensor structure would be challenging. Separate fabrication of the electronics chip and location apart from the sensor chip would be far less risky and, in spite of the problem of interconnection, probably more economical for a high-performance device.

Interaction between Sensors and Amplifiers

Analysis of the combined noise of a sensor and an amplifier is more complex than simply adding the two contributions. The simplest model of any value is shown in Fig. 3.4. The amplifier is modeled as an ideal gain stage and the equivalent voltage and current noise components associated with the amplifier are included as separate generators. Calculations based on this model normally reflect the actual system noise accurately. For the present purposes, all of the noise generators will be treated as if they are completely uncorrelated. Sometimes the processes are correlated but the errors introduced by unwarranted assumptions of zero correlation are usually small.

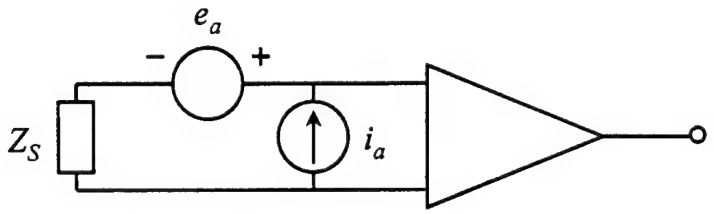

Figure 3.4. Noise model for an amplifier. The real amplifier is replaced by a noise-free gain stage and two noise generators – a voltage-noise source and a current-noise source. The real part of the sensor impedance, Z_S, also contributes Johnson noise.

As before, all noise values will be referred to the input of the amplifier. This is not necessary but it is conventional. (Ultimately, we would convert the raw noise values to noise-equivalent signal values.) There are three components of noise in this model. The first is the equivalent voltage noise of the amplifier, which is already in the form of voltage noise referred to the amplifier input:

$$e_1^2 = e_a^2 \tag{3.14}$$

The second results from the amplifier current noise flowing into the sensor impedance. In reality, this current noise flows into the parallel combination of the sensor impedance and the amplifier input impedance (and cable impedance, etc.) but this is not a fundamental complication.

The equivalent voltage noise is then

$$e_2^2 = i_a^2 |Z_S|^2 \tag{3.15}$$

The third is the intrinsic noise of the sensor, which is related by a Nyquist expression to the real part of the sensor's electrical impedance:

$$e_3^2 = 4 k_B T \, \mathrm{Re}[Z_S] \tag{3.16}$$

In principle, an accurate measurement of the sensor's electrical impedance will also reflect mechanical loss terms converted into their electrical equivalents so that the total noise is modeled correctly. Most often, though, analyses are based on incomplete models or measurements and it may be more convenient to treat the mechanical-thermal noise contribution separately after conversion to its equivalent electrical signal level. Some care should be used to avoid missing components or counting components twice (although less harm is done by the latter).

The total mean-square voltage noise is the sum of the individual mean-square values. By measuring the output noise of an amplifier connected to a wide range of resistors, the motivation for this two-component model for the amplifier noise becomes clear. Such a measurement (performed at 1000 Hz) is illustrated in Fig. 3.5. Resistors from 1 ohm to 1 gigaohm were connected across the input terminals of two different operational amplifier chips. The output noise was measured and then divided by the amplifier gain to produce input-equivalent values. The input-equivalent noise voltage was then plotted as a function of resistance value. In addition, the straight line of the mean-square Johnson noise associated with the resistance is plotted as a dotted line.

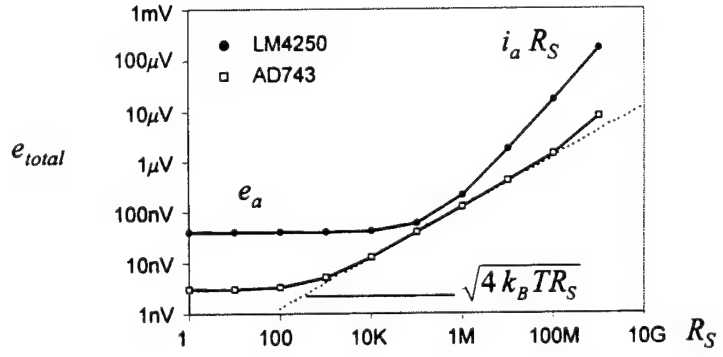

Figure 3.5. Total input-equivalent noise as a function of source resistance for two operational amplifiers. The low-noise AD743 has a significant range over which its noise is negligible compared to a resistive sensor. The dotted line represents the Johnson noise of R_S, which is independent of the amplifier characteristics.

Both of the curves have three sections. For small resistance, the curves are horizontal because they are dominated by the amplifier's voltage noise (e_1), which is not dependent on the source

resistance. For large resistance, the LM4250 curve has a slope of one because it is dominated by the amplifier's current noise flowing into the source resistance (e_2). The AD743 curve only shows the start of the transition to a slope of one. For some intermediate range of resistance, the curve either merges with or approaches the Johnson noise line (e_3).

The LM4250 is a micropower amplifier and, as such, is not a low-noise device. There is no region over which the resistor's intrinsic (Johnson) noise is accurately reproduced. However, the total noise approaches closely for a source resistance of about 250 kΩ, which is the effective noise resistance of the LM4250. The AD743, on the other hand, is a low-noise device and there is a substantial range of resistance for which the amplifier's contribution is small compared to the resistor's Johnson noise. The amplifier's contribution is less than the Johnson-noise contribution from about 400 ohms to about 400 MΩ. The equivalent noise resistance of the AD743 is about 400 kΩ and the normalized noise power is 0.001. Notice that, if the noise resistance is multiplied and divided by the normalized noise power factor, the bounds of the region for which the amplifier noise is less than the Johnson noise (i.e., $e_1^2 + e_1^2 < e_3^2$) are found.

For some amplifiers, the amplifier voltage and current noise components are correlated to some degree. Noise components add as mean-squares if uncorrelated and as amplitudes if fully correlated. Consequently, the worst error from a bad assumption regarding correlation is a factor of $\sqrt{2}$ in amplitude. That maximum error occurs when the e_1 and e_2 are equal. For a low-noise amplifier and a resistive sensor, the noise is dominated by the sensor's Johnson noise when e_1 and e_2 are equal (see Fig. 3.5) so the correlation assumption is irrelevant.

While we have examined the fundamental limits associated with e_a and i_a, it is useful to examine the ranges of e_a and i_a available in commercial operational amplifier chips. The values (at 1000 Hz) for an assortment of amplifiers[3] are shown in Fig. 3.6.

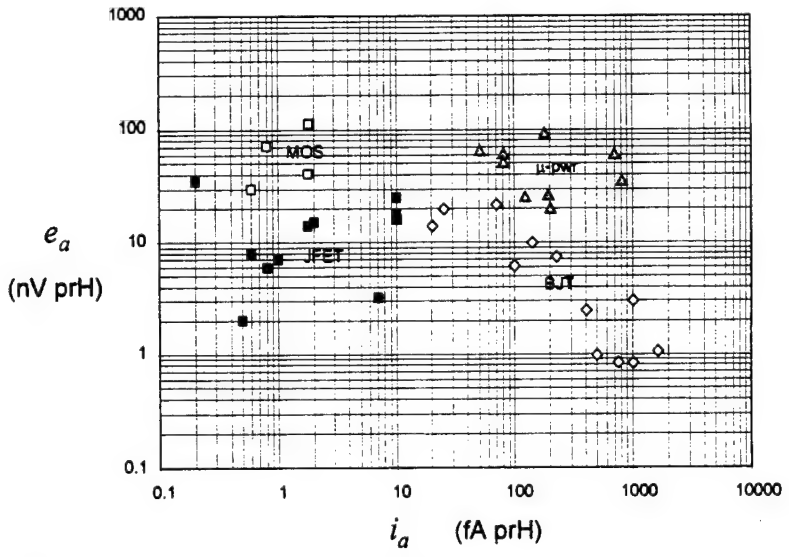

Figure 3.6. Scatter plot of the voltage and current noise values for several different operational amplifier chips. The solid squares are JFET devices, the open diamonds are bipolar (BJT) devices, the open squares are metal-oxide semiconductor (MOS) devices, and the triangles are micropower (μ-pwr) devices.

Of particular note is the grouping by transistor type. Those amplifiers with junction field-effect transistors in the input stage have the lowest current noise while the bipolar-input devices have the lowest voltage noise. Micropower chips are generally poor on both counts, while the metal-oxide semiconductor (MOS) devices have low current noise but very high voltage noise.

Notice that few devices are below 1 nanovolt per root hertz or 0.5 femtoamps per root hertz. A good bipolar device may have a voltage noise of one nanovolt per root hertz and a current noise of one picoamp per root hertz; a good JFET device may have a current noise of 1 femtoamp per root hertz and a voltage noise of 10 nanovolts per root hertz. In the design of low-voltage-noise systems, achieving one nanovolt per root hertz is possible but requires careful design while a level of ten nanovolts per root hertz is relatively easy. In the design of low-current-noise systems, achieving one femtoamp per root hertz is possible with considerable care while ten femtoamps per root hertz is straightforward.

If we plot the same assortment of devices using the normalized noise power and the equivalent noise resistance as coordinates (see Fig. 3.7), we see other associations.

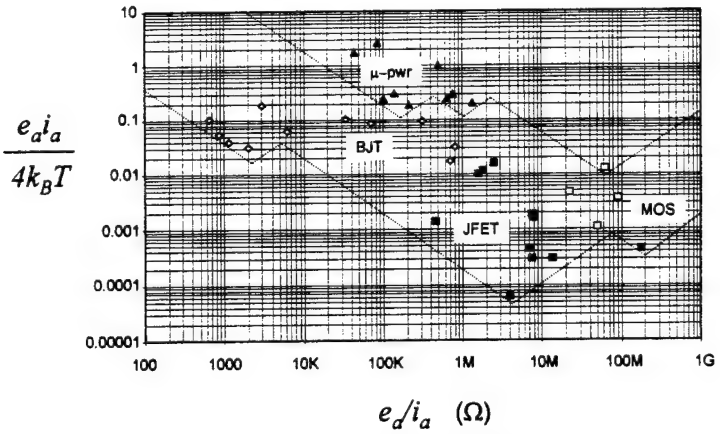

Figure 3.7. The same assortment of devices as in Fig. 3.6 with normalized noise power plotted as a function of effective noise resistance.

The JFET devices have, by far, the lowest normalized noise powers. While they cluster at higher values of equivalent noise resistance, these devices would be useful over the largest spans of source resistance. The BJT devices, on the other hand, are most useful for low-impedance sensors.

Knowing the ranges for e_n and i_n of available amplifiers permits intelligent design of a sensor/electronics interface. As long as the sensor impedance, Z_S, is known, all three of the basic noise components (Eqs. 3.14-3.16) can be calculated. The requirements for e_n and i_n can then be determined based on the desired system noise floor. A detailed example of this process is in Chapter 5 regarding the differential-capacitance accelerometer.

While the discussion above covers the important fundamentals of sensor-amplifier interaction, two other issues may need to be addressed. At very low frequency, additional noise with a power spectrum proportional to (frequency)$^{-1}$ can dominate. Such "1/f" noise components cannot generally be predicted from fundamentals but amplifier-chip manufacturers provide the expected levels for their products. For very low frequency systems, these 1/f components may dominate the noise floor.

The second non-fundamental issue is that of electromagnetic interference. While this can be extremely troublesome for some sensors, it is not a limiting issue for the accelerometer described below because of the inherent rejection of electromagnetic interference in the AC-drive configuration.

---

[1] F. Reif, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York, 1965.

[2] K. Lonngren, *Introduction to Physical Electronics*, Allyn and Bacon, Boston, 1988.

[3] Data taken from specification sheets from Analog Devices, Linear Technology, Siliconix, National Semiconductor, Maxim, and Comlinear. Various commercial products are mentioned in this report but none of these references should be construed as endorsements or recommendations.

Chapter 4

Structural Limits in Microfabrication

Microfabrication has enabled production of a wide variety of mechanical structures with dimensions in the range between micrometers and millimeters. While there are a number of successful fabrication techniques, each of these techniques has significant limitations that must be considered in sensor design. In this chapter, we will consider some of the limitations associated with microfabrication processes and microstructures. While we will concentrate on those issues concerning low-noise accelerometers, most of the discussion is applicable to microfabricated sensors in general.

Basic Scaling

Often, performance achieved on one scale can be extrapolated to another scale with some success. Such extrapolation, however, is critically dependent on an understanding of the fundamental processes involved. For example, a species of bug can swim underwater at up to 10 centimeters per second. The bug is 1 millimeter in length. Therefore, if we can emulate the propulsion system on a larger scale, we could produce an underwater vehicle 1 meter in length that would travel up to 100 meters per second! But, this sort of speculative "scaling" is fraught with misunderstanding. In the motion of submerged bodies, there are two principal forces – inertial forces and viscous forces. We can draw legitimate parallels between two systems if the ratio of inertial forces to viscous forces is the same for both. The nondimensional parameter expressing this ratio is the Reynolds number:

$$Re = \frac{\rho v L}{\eta} \tag{4.1}$$

where ρ is the fluid density (1000 kg/m^3 for water), v is the velocity of the body, L is the length (or diameter) of the body, and η is the fluid viscosity (0.001 Pa·s for water at 20°C). The Reynolds number for the bug in water is about 100. If the length were increased to 1 meter, the same Reynolds number would be achieved at a speed of 0.1 meters per second – a far less impressive performance!

One of the fundamental relationships associated with miniaturization of structures is the area-to-volume relationship. Some effects scale with volume (inertial forces, heat capacity, mass) and some scale with surface area (heat transfer, electrostatic force). If an object is reduced by a factor of two in all dimensions, its volume drops to one-eighth of the original volume while the surface area drops to one-quarter of the original area. Consequently, we expect surface phenomena to become increasingly important with miniaturization and, in general, this is true. The effects of area-to-volume ratio are particularly obvious in biological organisms. Single-celled organisms can transport food, oxygen, and waste products directly through their cell walls. Mammals have a much smaller surface-area-to-volume ratio so they must use area-enhanced organs such as lungs and intestines for these exchange processes. In microstructures, thermal

response is faster because heat capacity drops more quickly than heat transfer and electrostatic forces are more effective in countering inertial forces.

Surface attraction forces (van der Waals forces or surface adsorption, for example) become significant. Surface attraction creates the problem of "stiction" in which surfaces that are brought together do not separate. Consequently, release etching of surface-microfabricated structures must be done carefully. A liquid rinse draws two surfaces together through surface tension as the rinse liquid evaporates. The surface tension force can bring the surfaces into contact sufficiently intimate that they adhere. Even if the release is successful, if the surfaces are brought into contact during operation, they may become permanently attached.

In scaling, one must be exceedingly careful to identify important forces and parameters. When scaling is attempted over an order of magnitude or more it is easy to overlook phenomena. Forces negligible on one scale may dominate on another. For example, experience with conventional accelerometers suggests that the noise floor is almost always determined by the first-stage preamplifier. In Chapter 2 we saw that, on very small scales, molecular agitation may take over as the dominant mechanism. If the designer is unaware of the fluctuating forces associated with thermal agitation and simple extrapolation based on preamplifier limits is done, unachievable performance predictions may be made.

Microfabrication

In surface microfabrication, thin films of polysilicon, glasses, silicon nitride, and metals are deposited, patterned, and etched in a sequence designed to produce an essentially two-dimensional structure. The structure is held in place during fabrication by sacrificial layers, then a release etch removes the constraining layers and frees the moveable elements of the structure. Important features of surface microfabrication are the unidirectional build-up of structure, the two-dimensional aspect of the structural elements, and the accumulation of topographic irregularities from layer to layer.

In bulk microfabrication, the substrate is patterned and etched directly. These structures can be as thick as the substrate wafer (typically no thicker than 500 micrometers). Anisotropic chemical etches produce well defined but asymmetric structures by etching preferentially along crystal planes in the silicon. Other etching techniques are not constrained to follow crystal planes but high-aspect ratio etching of symmetric structural elements is still challenging. Double-sided processing is possible but requires careful back-to-front alignment of photolithography masks. To control the depth of chemical etches, various etch-stop mechanisms are employed. Timed etches are simple but lack repeatability. Silicon nitride resists common silicon etchants and is often used both to stop etches and to form beams or diaphragms. Ion implantation (particularly boron implantation) increases the resistance of silicon to common chemical etchants dramatically and is also used as an etch stop.

Structures of substantial complexity have been made through microfabrication by creative adaptation of these processes but it is important to realize that there are limits and these limits affect the ultimate performance of a device. For example, if low noise floor is important,

and all of the technologies were studied as a function of sensor size, we would see a pattern similar to that shown in Fig. 4.1.

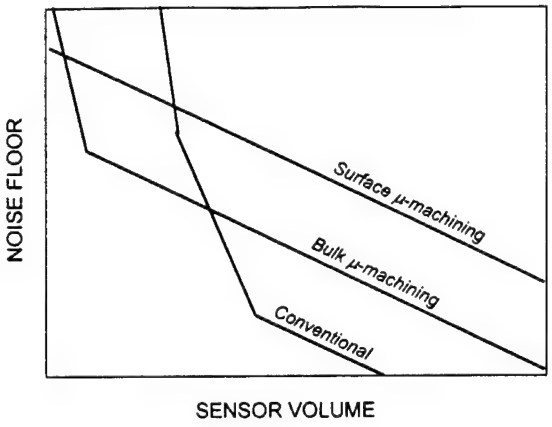

Figure 4.1. Qualitative comparison of potential for noise performance as a function of sensor volume for three processes. For example, surface microfabrication can produce the smallest structures but, on any scale for which other processes are feasible, the surface-microfabricated sensors would generally have higher noise floors.

On the centimeter scale, conventional machining techniques can be employed, many materials can be used, and many structural configurations are practical. With this design freedom, many alternatives are available for optimizing the performance of the device. Bulk microfabrication is more limited in materials and processes than conventional fabrication, so, for the same size scale, the limiting noise floor is higher. Surface microfabrication is further limited to very small proof masses so it would have the highest noise floor. However, as the size decreases, conventional machining becomes impractical. On the millimeter scale, bulk microfabrication has the most potential for low-noise accelerometers but, as the size drops even more, bulk microfabrication becomes impractical. On the scale of tens of micrometers, surface microfabrication becomes the process of choice although the price of extreme miniaturization is very poor noise performance. Predicting the achievable noise floor is not simply a matter of extrapolation of the performance achievable with conventional manufacturing; the constraints of more restrictive fabrication processes must be considered.

Micro-Arrays of Sensors

In principle, large numbers of micrometer-sized sensors can be made economically through batch fabrication. The suggestion is sometimes made to combine many such sensors so that the overall performance would exceed that of a single, larger sensor. While the performance gain is theoretically possible, it is worth examining this idea more closely.

If the ultimate limit to detectability (the thermal limit) is translated from Eq. 2.11 to the ratio of signal-to-noise power, then, the signal-to-noise ratio is proportional to the following combination of Q, moving mass, and resonance frequency:

$$SNR \rightarrow \frac{Qm}{f_0} \tag{4.2}$$

This expression argues for minimizing the resonance frequency, maximizing the Q, and maximizing the proof mass. (Of course, the desired bandwidth, dynamic range, and transient response impose practical restrictions on Q and f_0.)

Given this fundamental limit, we can consider improvements in performance with arrays of sensors. One of the touted advantages of microfabrication is the economy of batch fabrication. The fabrication costs are the same whether one or one hundred sensors are produced from a single wafer of silicon. (This argument is only plausible in the present context. If *individual* sensors are being made, the cost of packaging often overwhelms the basic chip cost.) If we are free to make many sensors, then we may be able to combine the outputs of many sensors. If the noise is not correlated from device to device but the signal is, then the signal *power* output of a combination of N devices would be N^2 times larger than for a single device and the noise *power* output would be N times larger. Therefore, the signal-to-noise power ratio is improved by the number of sensors, N:

$$SNR_N \rightarrow N * SNR \tag{4.3}$$

In microfabrication, however, there is considerable volume overhead in each single sensor (particularly for surface microfabrication). Consider an accelerometer structure with a design volume of 1 cm by 1 cm by a few millimeters. The proof mass for a bulk-microfabricated accelerometer might be 0.5 cm by 0.5 cm by 250 μm or 14 milligrams in silicon. By contrast, the proof mass of the surface-microfabricated Analog Devices ADXL100 series accelerometer is 0.1 microgram. More than 100,000 of these surface-microfabricated devices would have to be assembled to equal the mass of the bulk device. The total volume of the array would be much larger (and the power consumption would be much higher). Surface-microfabricated devices are very small but they use the substrate *thickness* very inefficiently. However, if relative noise performance is unimportant, and we just compare single sensor elements, then the economics of batch fabrication make the surface-microfabricated device an order of magnitude cheaper than the bulk device.

Residual Stress Management

Beyond simple scaling of area and volume, there are several other fundamental issues associated with the microfabrication process. One of these is residual stress. Many of the processes in microfabrication produce residual tensile or compressive stress in various layers or parts of structures. Residual stress can be problematic in conventional machining also, but the problem is aggravated by the small scales and intrusive processes of microfabrication.

Annealing processes can mitigate residual stress problems in some structures but they are generally ineffective when the stresses are caused by differences in material composition.

For example, to produce an etch-stop layer in silicon, boron atoms may be implanted to some depth in the silicon. The ordinary silicon would be etched away leaving a beam or diaphragm of boron-implanted silicon. This is an effective, controllable technique; however, the boron atoms are smaller than the silicon atoms. Consequently, the implantation produces a strong residual tensile stress (see Fig. 4.2). The residual stresses produced by typical levels of boron implantation range from 40 to 80 MPa. The tensile stress can be offset by also implanting large atoms – germanium atoms, for example. The residual stress can be reduced or even made compressive by sufficient implantation of the larger atom but the results are less repeatable.

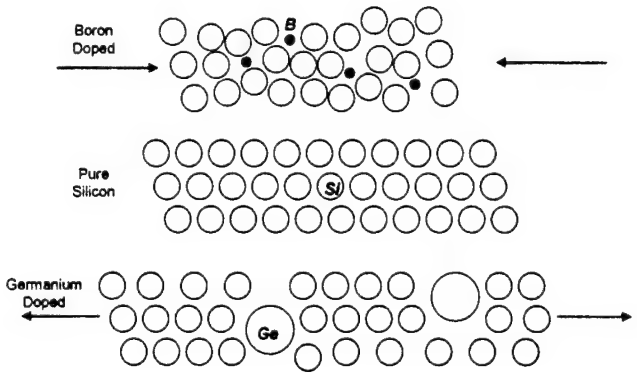

Figure 4.2. Effects of large and small atoms implanted in single-crystal silicon. Boron atoms are smaller than silicon atoms and produce residual tensile stress; germanium atoms are larger than silicon atoms and produce residual compressive stress. The arrows indicate the direction of dimensional change. The boron-doped silicon, for example, shrinks with respect to fixed supports so the residual stress is tensile.

For some structures, residual tensile stress is desirable. A thin membrane with tensile stress is flatter than without residual stress. A beam structure can be fabricated with residual tension so that subsequent compressive stress (introduced by differential thermal expansion, for example) is insufficient to cause buckling. However, modeling of the beam dynamics must consider this stress. A strong residual tensile stress may yield a structure that behaves more like a stretched string than an elastic beam. In addition, the consequences of the reaction stresses where the beam or membrane is attached to other parts of the structure need to be considered.

Other processes can leave residual compressive stress (thin-film deposition of polysilicon, for example). If the compressive stress is sufficiently large, beam elements can buckle (see Fig. 4.3). If the residual stress is large enough (either tensile or compressive), structural elements can fail. The critical stress, S_{crit}, for beam buckling is[1]

$$S_{crit} = \gamma \pi^2 \frac{E I}{L^2 A} \tag{4.4}$$

where E is the Young's modulus of the material (190 GPa for silicon), I is the moment of inertial of the beam cross-section, L is the length of the beam, and A is the cross-sectional area of the beam. The factor γ depends on the end conditions of the beam. If the ends can rotate (as if they were pinned), then $\gamma = 1$; if the ends are completely clamped, then $\gamma = 4$. For the end conditions achievable in microfabrication, γ would be somewhere between one and four.

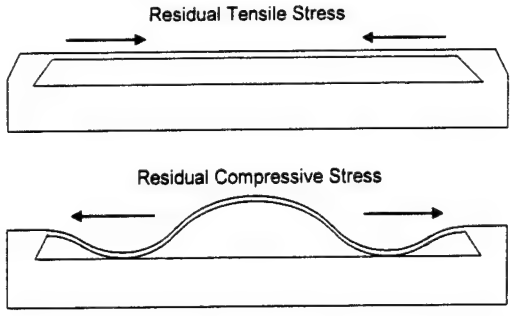

Figure 4.3. Residual stress in beam or membrane structures. Tensile stress produces flat structures with high stress at points of attachment. Compressive stress can cause buckling of structures if sufficiently high.

The parameter γ can also be viewed as the dimensionless parameter for buckling. For beams of rectangular cross-section,

$$I = \frac{wt^3}{12} \quad ; \quad A = wt \tag{4.5}$$

where w and t are the beam width and thickness, respectively. Then, the buckling parameter is

$$\gamma = \frac{12}{\pi^2}\frac{S_{crit}}{E}\left[\frac{L}{t}\right]^2 \tag{4.6}$$

Design scaling for surface microfabrication tends to reduce thickness, t, considerably more then length, L. Consequently, buckling is more likely to occur in these designs. With a residual stress of 60 MPa and the Young's modulus for silicon (190 GPa), the buckling parameter would be greater than one for length-to-thickness ratios greater than about 50. A 2-micrometer thick beam could be no greater than 100 micrometers in length or it may buckle.

Sufficient residual tensile stress can produce a structure in which the tension forces rather than the bending moments control the deflection. The relative importance of these two mechanisms can be determined from their contributions to the strain energy in the beam. The bending strain energy per unit length, dU_B, is given by

$$dU_B = \frac{1}{2} EI\left(\frac{d^2y}{dx^2}\right)^2 dx \tag{4.7}$$

where y is the transverse deflection as a function of x, the length coordinate. The tensile strain energy per unit length, dU_T, is given by the work against the residual stress, S, or

$$dU_T = \frac{1}{2} S A \left(\frac{dy}{dx}\right)^2 dx \tag{4.8}$$

For a beam that is fixed at both ends ($x = \pm L/2$) and has a deflection at the center of y_0, the deflection function (without residual stress[2]) is

$$y = \frac{y_0}{2}\left[\cos(2\pi x/L) + 1\right] \tag{4.9}$$

If we integrate the required derivatives of Eq. 4.9 over the length, L, of the beam and write the ratio of the tensile strain energy to the bending strain energy, we have

$$\frac{U_T}{U_B} = \frac{12}{4\pi^2}\frac{S}{E}\left[\frac{L}{t}\right]^2 \tag{4.10}$$

If this ratio is larger than one, then the structure is more like a stretched string than an elastic beam and calculations of dynamic behavior and static deflections should be made accordingly. Notice the similarity to Eq. 4.6.

Layered processing in microfabrication leads to gradients in residual stress. Gradients in stress can cause warping of structures. Especially in surface microfabrication, in which several layers of different material are deposited in sequence, vertical gradients in residual stress are common. In order to evaluate the effects of a stress gradient, consider the distribution of fiber stress with distance, z, away from the neutral axis in an ordinary beam with some deflection. The stress is

$$S(z) = \frac{Mz}{I} \tag{4.11}$$

where M is the bending moment. The bending moment is related to the deflection function, $y(x)$, where x is the coordinate along the length of the beam, by

$$M = EI\frac{d^2y}{dx^2} \tag{4.12}$$

The same deflection would be produced by imposing the same stress gradient as in Eq. 4.11 or

$$S' \equiv \frac{dS}{dz} = \frac{M}{I} = E\frac{d^2y}{dx^2} \tag{4.13}$$

If the stress gradient, S', were constant and the beam were clamped at $x = 0$, then the deflection function would be

$$y = \frac{1}{2}\frac{S'}{E}x^2 \tag{4.14}$$

If the stress gradient is written as a ΔS over the beam thickness, then the ratio of deflection to thickness at the end of the beam ($x = L$) would be

$$\frac{y_L}{t} = \frac{1}{2}\frac{\Delta S}{E}\left[\frac{L}{t}\right]^2 \tag{4.15}$$

Except for the constant factor, this expression has the same form as that for the buckling parameter (Eq. 4.6) with similar consequences for microfabrication of large length-to-thickness ratio beams. If the structures are not designed to be sufficiently stiff, the gradient-induced warp may be unacceptably large (see Fig. 4.4).

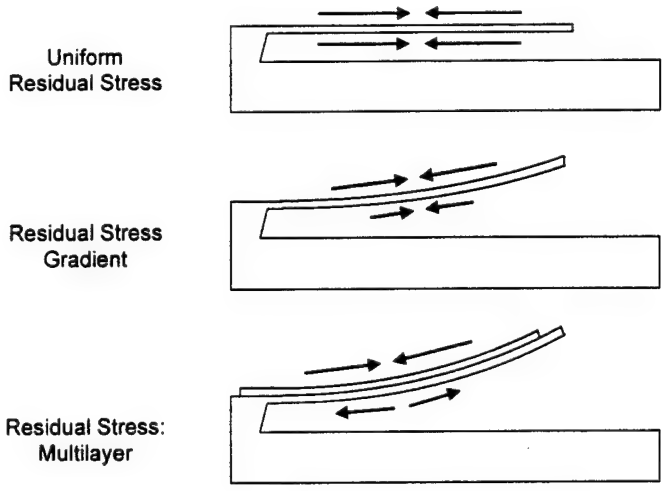

Figure 4.4. Warping of structures caused by gradients in residual stress. In the lower structure, the gradient is caused by adjacent layers of two different materials each having a different residual stress. In the middle structure, the gradient is caused by different processing on one side of a structure than on the other.

The structure shown in Fig. 4.5 is an extreme case. Here, the warping was sufficient to completely misalign the two elements of an electrostatic comb structure.

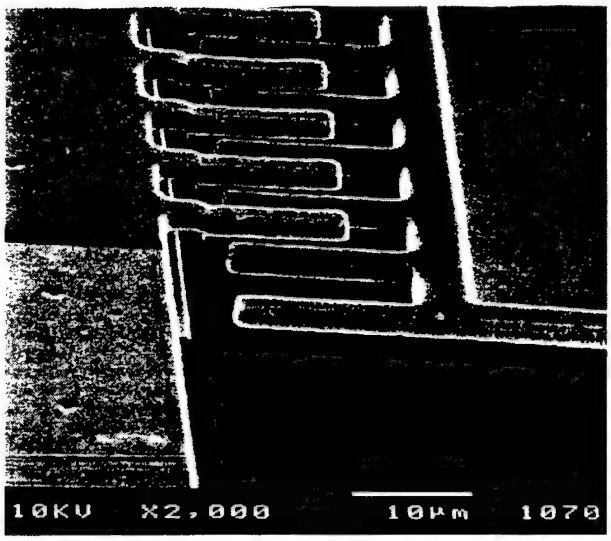

Figure 4.5. Surface-microfabricated comb structure. One side of the comb is warped so much that the fingers are completely misaligned.

While it is difficult to reduce residual stresses to levels that have no significant influence on the structure, it is, at least in principle, possible to design for management of residual stresses. The basic design principle involves structural symmetry. If the structure can be designed so that it is symmetric about a plane and that plane is perpendicular to the direction of maximum stress gradients, then the gradients will balance. This is illustrated in Fig. 4.6. The upper structure is representative of a typical microfabrication process. The gradient in residual stress is predominantly vertical and the structural element responds to the imbalance in stress by warping.

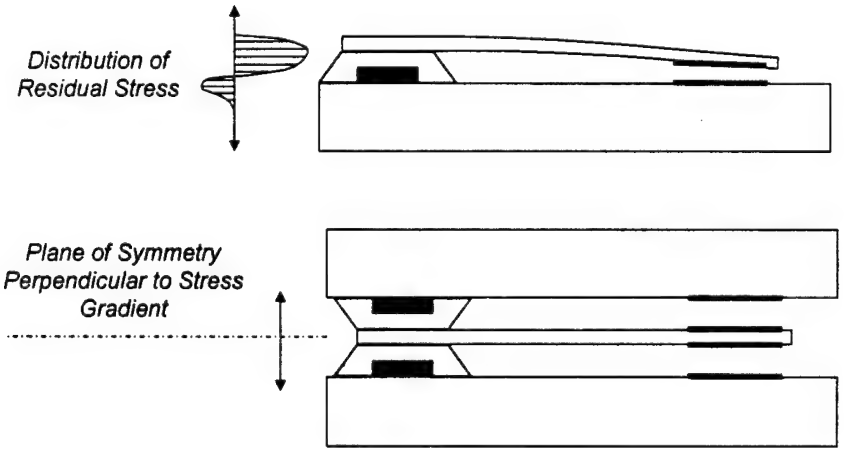

Figure 4.6. Balancing of residual stress gradients to prevent distortion of structures. Unbalanced stresses produce warping in the upper structure. The same fabrication stresses are introduced in the lower structure but they are introduced symmetrically. The balanced gradients minimize warping.

In the lower structure, the structure has been designed so that it is symmetric about the moving structural element (the cantilever). The *level* of residual stresses is the same as in the upper structure but now any stresses below the cantilever are balanced by the same stresses above the cantilever. The gradients are as large as before but they are balanced symmetrically about the critical structural element. This principle of symmetry also minimizes the effects of temperature changes in the structure for the same reason – the thermally induced stresses balance.

Unfortunately, such symmetric structures are difficult to produce by microfabrication. One option is to fabricate two half-structures and bond them together back-to-back, but this requires a difficult alignment step.

Fluid Dynamics

Miniaturization of structures also leads to interesting problems in gas flow. Beyond the noise produced by molecular agitation, viscous damping of moving surfaces can be substantial. We will consider two cases. In one case (Fig. 4.7), the moving surface moves toward and away from the stationary surface so that the gas in between is alternately squeezed out toward the periphery and pulled in toward the center. In the other case (Fig. 4.8), the moving surface oscillates parallel to the fixed surface so there is no change in gas volume between the surfaces but there is still a velocity gradient.

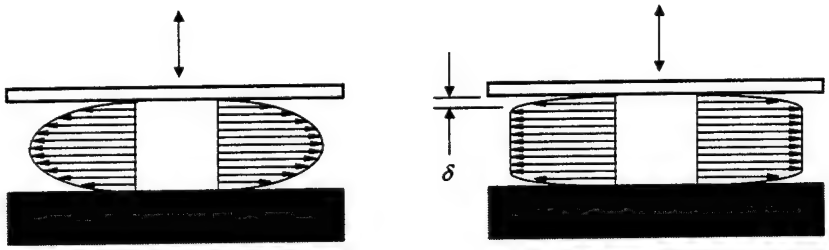

Figure 4.7. Velocity profiles for gas flow between plates. The upper plate is oscillating toward and away from the lower plate. (Flow profiles are shown for downward motion.) On the left, the viscous penetration depth, δ, is larger than the plate spacing and the profile is parabolic. On the right, the penetration depth is much less than the plate spacing and a distinct boundary layer is formed next to each surface.

An important scale distance in these gas-flow problems is the viscous penetration depth[3] (or viscous boundary-layer thickness), δ:

$$\delta^2 = \frac{2\eta}{\omega\rho} \tag{4.16}$$

where ω is the angular frequency of the oscillatory motion, η is the viscosity of the gas ($18 \cdot 10^{-6}$ Pa·s for air at 20° C) and ρ is the density of the gas (1.2 kg/m^3 for air at atmospheric pressure and 20° C). If the penetration depth is much less than the spacing between surfaces, then a distinct boundary layer forms in which the bulk of the velocity gradient resides. Outside the boundary

layer, the flow speed is roughly constant. If the penetration depth is larger than the spacing (the more common instance for microfabricated devices), then the flow profile is roughly parabolic with a maximum halfway between the surfaces. In either case, there is a gradient in the velocity and, consequently, viscous drag. The discussion below is limited to the case in which the penetration depth is larger than the spacing.

For the case of two parallel, circular disks moving toward and away from each other, the mechanical resistance (the "squeeze-film" resistance) is[4],

$$R_{mech} = \frac{3\eta A^2}{2\pi h^3} \tag{4.17}$$

where A is the surface area of one of the disks, and h is the equilibrium spacing between them. Notice the dependence on the inverse third power of the spacing. This component of damping increases rapidly with miniaturization.

Often, one of the surfaces is perforated to reduce the damping. The component of resistance related to the flow through the perforations is given approximately by[5]

$$R_{mech} = \frac{12\eta B^2}{N\pi h^3} G(B) \tag{4.18}$$

where N is the total number of holes in the plate, B is the fraction of open area, and the function, G, is

$$G(B) = \frac{1}{8}\left(4B - B^2 - 2\ln B - 3\right) \tag{4.19}$$

The total mechanical resistance for a perforated-plate structure is the parallel combination of the resistances given in Eqs. 4.17 and 4.18.

The penetration depth is also a suitable scale length for transverse motion (Fig. 4.8). The important case for microfabricated structures is the case in which the penetration depth is larger than the plate spacing and the gas velocity is linear between the two surfaces. In that case, the velocity gradient is constant and equal to the plate speed divided by the plate spacing.

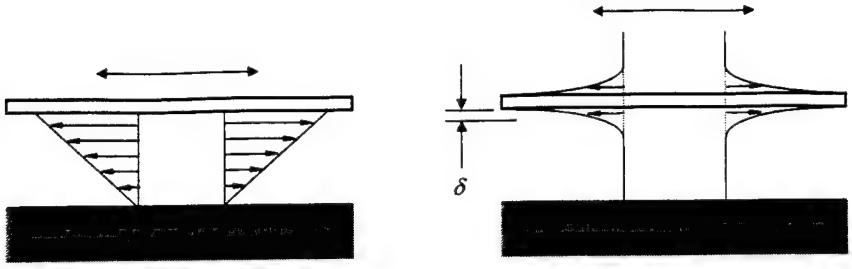

Figure 4.8. Velocity profiles for gas flow between plates in which the upper plate is oscillating parallel to the lower plate. (Flow profiles for both right and left motion are shown.) The viscous penetration depth, δ, is larger than the plate spacing in the diagram on the left and the flow profile is linear. On the right, the penetration depth is much less than the plate spacing and a distinct boundary layer forms near the moving plate. The flow profiles on the right are only schematic; the actual profile is more complicated.

The force caused by viscosity is

$$F_\eta = \eta A \frac{dv}{dz} = \eta A \frac{v_{plate}}{h} \tag{4.20}$$

so the mechanical resistance is

$$R_{mech} = \frac{F_\eta}{v_{plate}} = \frac{\eta A}{h} \tag{4.21}$$

As was true for the fluctuating force that results from molecular collisions, the viscous drag function also changes when the molecular mean-free-path approaches or exceeds the dimensions of the channel that confines the flow. For flow through a rectangular tube (Fig. 4.9) with height much less than either width or length, the ratio of pressure drop to flow speed is approximately constant as long as the mean-free-path is much less than the height of the tube.

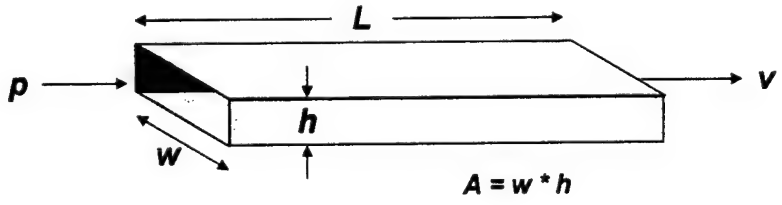

Figure 4.9. Geometry of a rectangular tube. Behavior of gas-flow through the tube is determined in part by the relationship between the molecular mean-free-path and the smallest dimension.

Once the mean-free-path reaches or exceeds the duct height, the pressure drop decreases for the same flow speed (Fig. 4.10). (Note: the speed here is particle speed not mass flow rate.) This

condition can occur for decreased pressure or for sufficiently small spacing between surfaces. There is a ten-percent effect at atmospheric pressure for a spacing of 6 micrometers.

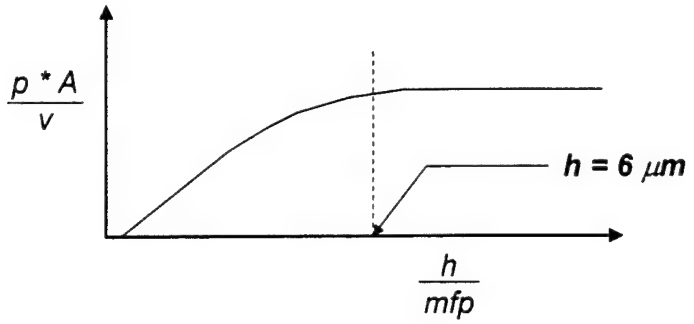

Figure 4.10. Normalized pressure drop through tube in Fig. 4.9. as a function of the ratio of tube height, *h*, to mean-free-path, *mfp*. The region of free-molecular flow is to the left and ordinary viscous flow is to the right. Compare this graph to the graph in Fig. 2.6.

In principle, the proof mass of an accelerometer can be operated in a vacuum. The Q would then be determined by mechanical damping in the structure and could be very large. Resonant sensors often employ this strategy for improved response to signal and some accelerometers are designed so that the input acceleration changes the resonance frequency of such a structure. Thermal agitation can still limit performance but the processes of control and detection of resonance need to be considered also. The more common commercial products employ gas damping, so knowledge of the fluid-dynamic damping is critical for prediction of the fundamental noise limits.

Increasing the Q decreases the equivalent acceleration noise, but, in the simple accelerometer, there are practical limits on the Q. If the Q is too high, then there will be a very large mechanical response in the structure at the resonance frequency. This can push the sensor into nonlinear operation or create dynamic range problems in the electronics interface. Sensors are often designed to be critically damped ($Q = 0.5$) so that the transient response is well behaved but Q's of 10 to 20 are practical. For Q's much larger than 10, the sensor may need to be operated with feedback in order to control the response around the resonance frequency. This adds additional complexity and the potential for instability. Feedback control is often implemented by electrostatic force in microstructures so it is worth examining these electrostatic forces in more detail.

Electrostatic Forces

Electrostatic forces scale favorably with miniaturization, so electrostatic actuation is often used in microstructures. For the simple structure shown in Fig. 4.11, the electrostatic force is

$$F_e = \frac{qV^2 A}{h^2} \tag{4.22}$$

where q is the charge on an electron, V is the voltage between the plates, A is the area of either plate, and h is the plate spacing.

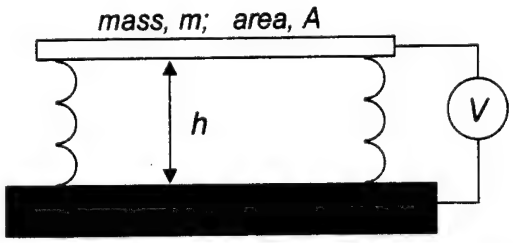

Figure 4.11. Simple model for electrostatic actuation. Two electrodes are separated by a spring system and some voltage is applied. The force is attractive and nonlinear. As the voltage is increased, the upper surface moves toward the lower surface. Beyond some point, the system is unstable and the surfaces "snap" together.

If the moving plate has mass, m, then the inertial force is mass times acceleration or

$$F_i = ma \tag{4.23}$$

The ratio of electrostatic force to inertial force is then

$$\frac{F_e}{F_i} = \frac{qV^2 A}{mah^2} \tag{4.24}$$

The moving mass, m, and the spacing, h, both decrease with miniaturization, which increases the dominance of electrostatic force over inertial force. Hence, electrostatic actuation is effective for microstructures.

However, stability under electrostatic actuation is an important consideration. The deflection is not linear with applied voltage. Beyond some critical displacement, the system is unstable and the surfaces snap together.

This so-called "snap-down" instability is normally analyzed for a linear spring[6]. For membranes and for beams fixed at both ends, the effective spring constant is often not linear for

large deflection so it is useful to consider the first non-linear correction, a cubic hardening. If the deflection in Fig. 4.11 is positive upward, then the restoring force for this nonlinear spring is

$$F_{spr} = -k_0 x - k_1 x^3 \tag{4.25}$$

The corresponding elastic potential energy is

$$U_k = \frac{1}{2} k_0 x^2 + \frac{1}{4} k_1 x^4 = \frac{1}{2} k_0 h^2 \mu^2 \left(1 + \phi \mu^2\right) \tag{4.26}$$

where the following normalized quantities have been introduced:

$$\mu \equiv \frac{x}{h} \qquad \text{and} \qquad \phi \equiv \frac{k_1 h^2}{2 k_0} \tag{4.27}$$

The electrostatic potential energy is

$$U_e = -\frac{1}{2} C(x) V^2 = -\frac{1}{2} C_0 V^2 \frac{1}{1 + \mu} \tag{4.28}$$

where C_0 is the capacitance for zero displacement. The total potential is the sum of U_k and U_e. If we normalize this sum by ½ $k_0 h^2$ and denote the normalized potential by U, we have

$$\mathsf{U} \equiv \frac{U_k + U_e}{\left(k_0 h^2 / 2\right)} = \mu^2 \left(1 + \phi \mu^2\right) - \frac{\gamma}{1 + \mu} \tag{4.29}$$

where

$$\gamma \equiv \frac{C_0 V^2}{k_0 h^2} \tag{4.30}$$

The equilibrium deflection occurs at the minimum of the potential or

$$\frac{d\mathsf{U}}{d\mu} = 0 \tag{4.31}$$

For a stable equilibrium, the second derivative is positive. The onset of instability is given by the vanishing of the second derivative:

$$\frac{d^2\mathsf{U}}{d\mu^2} = 0 \tag{4.32}$$

Substituting Eq. 4.29 for U into Eqs. 4.31 and 4.32 and eliminating γ between the two resulting equations gives the following equation for the point of instability:

$$10\phi\mu^3 + 6\phi\mu^2 + 3\mu + 1 = 0 \tag{4.33}$$

If the spring is linear, then k_1 and ϕ are both zero and the solution of Eq. 4.33 is

$$\mu_{crit} = -1/3 \tag{4.34}$$

This is the classic snap-down solution. For a linear spring, if the voltage is increased to the point at which the deflection is one third of the zero-voltage separation, the plate will snap down onto the fixed plate. If the spring has no linear component but is only cubic, then k_0 is zero and ϕ approaches infinity. The solution of Eq. 4.33 is then

$$\mu_{crit} = -3/5 \tag{4.35}$$

In this case, the voltage can be increased until the deflection is three-fifths of the zero-voltage separation before the snap-down instability. For the general linear-plus-cubic spring, the critical deflection is between 1/3 and 3/5.

This instability is important not only for the design of electrostatic actuation systems but also in the design of DC-bias capacitive sensors. It is, in fact, one of the important design considerations in condenser microphone design. Small gaps and high voltage both result in higher responsivity but require operation closer to the snap-down instability.

Three-Dimensional Effects

In microfabrication, the complexity of structures is limited primarily by the resolution and registry of photolithographic patterning and etching. Consequently, interdigitated "comb" structures that would be impractical to produce by conventional means are often employed for electrostatic actuation or capacitive sensing in surface-microfabricated structures. The use of photolithography and thin-film deposition tends to promote two-dimensional thinking in structure design. In reality, three-dimensional effects are important.

One three-dimensional effect is accumulation of topography. In microfabrication, the initial substrate surface is flat. As layers are sequentially deposited, patterned, and etched, new layers assume the topography of the layers beneath. A layer that covers only part of the substrate produces, in the next layer, a low mesa. If the next layer only partially covers the mesa, the following layer will show compound variations in "elevation" resulting from the overlapping areas.

A simple example of this is shown in Fig. 4.12. This is a scanning electron micrograph of a small section of an interdigitated comb structure. The moving plate is toward the top of the figure while the fixed plate is toward the bottom. A metal layer was deposited on the substrate

below the moving plate. The topography of the metal layer is reflected in the topography of the moving plate by the downward "jog" in each finger that appears just beyond the plate-finger boundary. The structure shown is quite simple. In more complex structures, the deviation from flatness is much more pronounced.

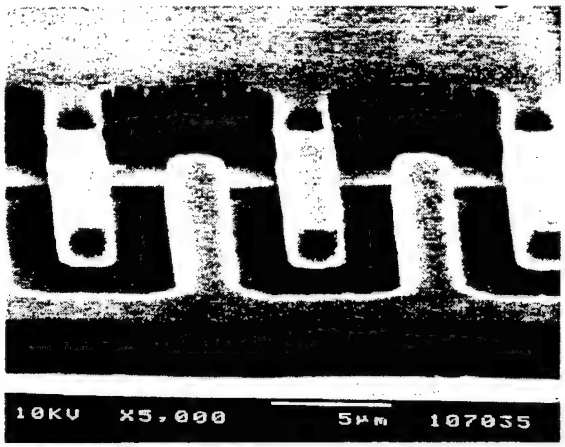

Figure 4.12. Interdigitated comb structure. The upper plate is attached to a compliant suspension. The layer underneath the upper plate is a metal electrode. The "step" in the upper fingers where they join the plate is a result of the electrode layer's topography being transferred to subsequent layers. This structure was designed to study the out-of-plane motion produced by electrostatic comb structures.

Another three-dimensional effect is particularly important in electrostatic comb actuators. When a comb is used as an actuator, a voltage is applied to some subset of the fingers so that the electrostatic attraction causes motion in the plane of the fingers. Because the gap between the comb structure and the substrate is small and the substrate is maintained at some potential, the substrate affects the electric field. Since the substrate is only on one side of the comb structure, the field is asymmetric and an out-of-plane force is produced. Not only does the moving element translate in the plane of the fingers; it also translates toward or away from the substrate. This complicates the response to control inputs and can degrade the performance of sensors designed for highly stable, linear output.

Even if the electrostatic field was symmetric, there is still a restoring force for out-of-plane motion. A symmetric field in a comb drive produces an approximately linear restoring force for in-plane motion and an approximately quadratic restoring force for out-of-plane motion. This results in parametric excitation of the moving element. If there is an out-of-plane resonance of the moving element at twice the frequency at which the comb is driven, large-amplitude out-of-plane motion can be produced.

When designing microstructures, although the geometry may be predominantly two-dimensional, three-dimensional effects must be considered. Both topography accumulation and electrostatic field instability and asymmetry can lead to unexpected fabrication or performance problems.

[1] W. Young, Roark's *Formulas for Stress and Strain*, 6th ed., McGraw-Hill, NY, 1989.

[2] If tensile stress is important, then this is not the proper deflection function. Energy methods have been used here so that approximate deflection functions are adequate for reasonable analysis.

[3] L. Landau and E. Lifshitz, *Fluid Mechanics*, Pergamon, NY, 1959, §24.

[4] J. Starr, "Squeeze-film damping in solid-state accelerometers," in *Solid-State Sensor and Actuator Workshop Technical Digest*, IEEE, NY, 1990.

[5] J. Bergqvist, F. Rudolf, J. Maisana, F. Parodi, and M. Rossi, "A silicon condensor microphone with a highly perforated backplate," in *1991 International Conference on Solid-State Sensors and Actuators*, Digest of Technical Papers, IEEE, 1991.

[6] A. Pippard, *Response and Stability*, Cambridge University Press, NY, 1985.

Chapter 5

The Differential-Capacitance Hydrophone Element

One of the most successful commercial MEMS accelerometers is the differential-capacitance accelerometer. The differential-capacitance structure is straightforward to fabricate by either surface or bulk micromachining. Surface micromachined structures are very small and often integrated on the same substrate as the drive and detection electronics. However, these structures have so little mass that it is challenging to reach even a milli-g per root hertz noise floor. Bulk microfabricated structures provide the large masses needed to reach micro-g levels or below but this potential is generally not exploited. The commercial market for such low-noise devices is small so incentive for development is weak. In this chapter, we will examine capacitive sensing from basic issues through a full-bridge, AC-drive sensor suitable for use in the directional hydrophone.

Capacitive Sensing

One of the simplest configurations (Fig. 5.1) for capacitive sensors involves applying a DC supply voltage, V_{SUPPLY}, through a large resistor, R_{SUPPLY}. This is representative of most commercial condenser microphones.

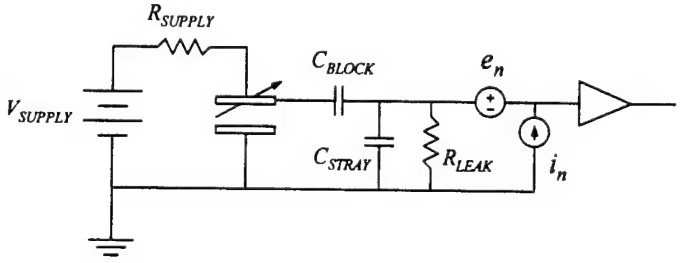

Figure 5.1. DC-drive capacitive sensor and preamplifier. The voltage on the moving plate is a DC voltage plus a voltage that varies with the same frequency as the excitation. Also shown are the equivalent noise generators, e_n and i_n, associated with the preamplifier, the DC-blocking capacitor, C_{BLOCK}, the stray capacitance, C_{STRAY}, and the leakage resistor, R_{LEAK}.

If the supply resistor is large, then, for plate motion at signal frequencies, the charge on the plates is constant. As the capacitor plate spacing changes, the capacitance and, therefore, the voltage change. In-band noise from the DC supply is filtered by the supply resistor but longer time scale changes in DC voltage directly affect the scale factor. The blocking capacitor, C_{BLOCK}, is required to prevent the large DC voltage from appearing at the preamplifier input. The capacitor, C_{STRAY}, represents "stray" capacitance associated with interconnections and the input of the amplifier. The resistor, R_{LEAK}, is required to provide a path for leakage currents from the amplifier to ground. The generators, e_n and i_n, are the equivalent noise generators of the preamplifier.

Perhaps the most serious problem with this configuration is the large sense-element impedance at low frequency. The sense element is a capacitor: the magnitude of its impedance is proportional to the reciprocal of frequency. The current noise of the amplifier flows into this impedance and generates an equivalent voltage noise, the amplitude of which is inversely proportional to frequency. This means that the preamplifier current noise usually dominates the overall noise floor at the low-frequency end of the signal band. Because the sense element's impedance changes over the band of interest, it can be challenging to design an appropriate electrical interface.

Many of the problems inherent in the DC sensor can be reduced or eliminated by AC-drive. While adding significant electrical complexity, this complexity can be justified for high-performance sensors. The configuration shown in Fig. 5.2 consists of a single differential element driven symmetrically at a frequency well above the band of interest. A high-frequency sine or square wave is applied to one fixed plate and its inverted replica is applied to the other fixed plate.

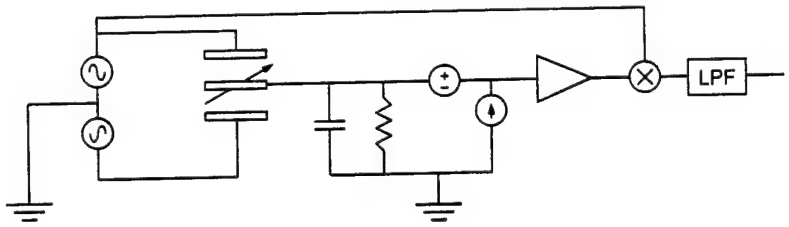

Figure 5.2. Basic AC-drive, differential-capacitance structure with first-stage preamplifier and detection electronics. The fixed plates are driven by out-of-phase sine waves at a frequency much higher than the expected signal excitation. The voltage on the moving plate is a high-frequency signal whose amplitude is proportional to the displacement of the moving plate from the center. The modulated signal is detected by multiplication and low-pass filtering (LPF).

If the moveable plate is centered, the sense signal is zero; if the plate is offset, the sense signal is a signal at the drive frequency with an amplitude proportional to the plate offset. If the acceleration input is sinusoidal at the frequency, ω_s, and the AC-drive signal to the plates is sinusoidal at ω_0, (with $\omega_0 >> \omega_s$) then the signal from the center plate is

$$e_{plate} = A\cos(\omega_s t)\cos(\omega_0 t) = \frac{A}{2}\left[\cos(\omega_0 + \omega_s) + \cos(\omega_0 - \omega_s)\right] \qquad (5.1)$$

In effect, the acceleration signal has been translated up in frequency to the vicinity of ω_0. After the sense signal passes through a preamplifier, the amplitude is extracted by multiplying the sense signal with the drive signal[1]. The output of the multiplier is

$$e_{mult} = A\cos(\omega_s t)\cos^2(\omega_0 t) = \frac{A}{2}\cos(\omega_s t)\left[1 + \cos(2\omega_0 t)\right] \qquad (5.2)$$

which consists of a demodulated (baseband) signal and a signal at twice the drive frequency. A low-pass filter removes the component at twice the drive frequency leaving the desired signal. Notice that, in contrast to Fig. 5.1, there is no blocking capacitor between the moving plate and the preamplifier. The plate signal has no DC component so no blocking capacitor is necessary. In addition, a static displacement of the plate produces a constant-amplitude, high-frequency signal so the system does respond down to zero frequency (or, equivalently, to static accelerations like that due to gravity).

The version in Fig. 5.3 is the configuration used in this investigation. Here, two differential elements are used to form a full bridge. This provides some compensation for variations in the static offset of the individual elements and provides some rejection of in-band drive-signal noise.

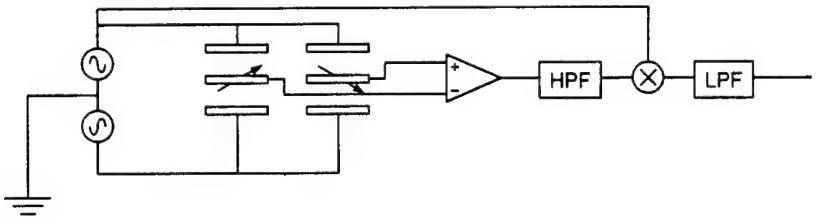

Figure 5.3. Full-bridge differential-capacitance configuration. The equivalent noise generators have been omitted for clarity. The sense signal is amplified differentially then high-pass filtered (HPF) to remove low-frequency interference. The detection process is identical to that in Fig. 5.2.

AC drive offers several significant advantages. First, the relevant impedance of the sense element (from the standpoint of self-noise) is the impedance at the drive frequency not the impedance in the signal band. By using a sufficiently high drive frequency, the region of noise dominated by preamplifier current noise (and 1/f components) can be avoided (see Fig. 5.4).

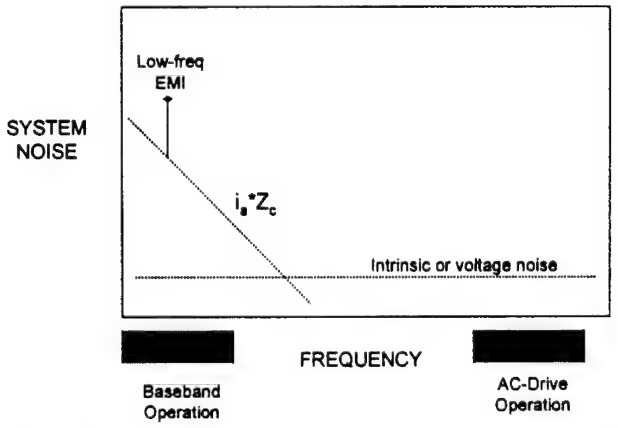

Figure 5.4. Comparison of baseband operation to AC-drive operation. Low-frequency baseband signals are in the same frequency range as electromagnetic interference (EMI) and the region in which the amplifier current noise produces the dominant component of noise. By translating the signal up in frequency, both of these problems can be avoided.

Experimental sensors have used drive frequencies in the gigahertz range but it is far more common to use drive frequencies below a megahertz. The smaller the basic element capacitance is, the higher the drive frequency to maintain the same sensor impedance. However, the drive frequency might be chosen to be a standard frequency used in some other commercial application (e.g., 455 kHz, 10.7 MHz, 900 MHz) in order to take advantage of low-cost, off-the-shelf electronics.

Another advantage that is rarely exploited is the inherent resistance to electromagnetic interference (EMI). Much of the troublesome EMI is in the range of 10 to 400 Hz and is coupled additively into high-impedance front-end interconnects and electronics. Eq. 5.1 becomes

$$e_{plate} = A\cos(\omega_s t)\cos(\omega_0 t) + B_{EMI} \tag{5.3}$$

A high-pass filter can be inserted (see Fig. 5.3) to reduce the low-frequency EMI, B_{EMI}, since the signal of interest has been translated to much higher frequencies.

The output of the multiplier is,

$$e_{mult} = \frac{A}{2}\cos(\omega_s t)[1 + \cos(2\omega_0 t)] + B'_{EMI}\cos(\omega_0 t) \tag{5.4}$$

The EMI remaining after the initial high-pass filter, B'_{EMI}, is translated by the multiplier to the vicinity of the high-frequency drive signal and is further rejected by the low-pass filter that follows the multiplier. Consequently, the plate signal prior to detection can be transmitted through "dirty" environments, then high-pass filtered and detected once the signal reaches a cleaner environment.

One issue that must be considered when using high-frequency drive is crosstalk. If there is direct electromagnetic coupling between the drive-signal interconnects and the center-plate-to-preamplifier connection, then a component of the output will be associated with this false signal. If the level of crosstalk is stable, then this is not in itself a problem. However, if the crosstalk is too high, the gain of the preamplifier may be small enough that the noise associated with the multiplier limits the overall performance. Careful design is required to minimize crosstalk, which is one of the more frequent causes of disappointing performance in AC-drive sensors.

Basic Configuration

The differential elements for this investigation were purchased from a commercial manufacturer[2] of accelerometers. They are bulk-micromachined structures made from single-crystal silicon. The structures are shown in cross-section in Fig. 5.5. The moving mass is about 2.5 milligrams, the one-sided capacitance (including stray capacitance) is about 35 pF, and the overall dimensions of the chip are 4 x 2.5 x 1 mm. Notice that, while the differential-capacitance structure can be made symmetric to minimize structural offsets (see Chapter 4), the proof-mass layer of this chip is not symmetric about the hinge plane. With chemical etches, it is simpler to produce the asymmetric center layer shown in the figure.

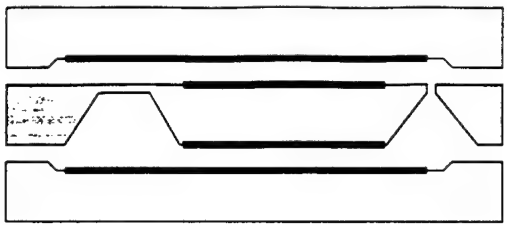

Figure 5.5. Cross-section through the three-layer MEMS structure comprising the differential-capacitance element. The thick black lines represent the electrodes. The two electrodes on the center, moving element are connected. The diagram is not to scale.

For this investigation, a single accelerometer element was made by bonding two differential elements back-to-back (Fig. 5.6). This enabled electrical connection as a full bridge.

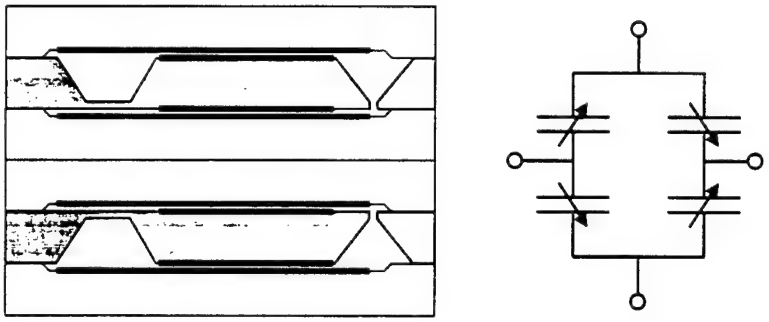

Figure 5.6. Two MEMS elements are bonded back-to-back to form a full-bridge structure.

The electrodes were connected so that an applied acceleration produced a differential signal. This could have been done without reversing one of the structures; however, the reversal permits some cancellation of the offset inherent in the asymmetric proof-mass structure. When a single element is oriented so that the sense axis is horizontal (*i.e.*, perpendicular to the direction of the local gravity vector), the center plate is not centered between the fixed plates: there is some zero-g offset or bias. While there was a substantial variation in this offset from chip to chip, there was a distinct tendency toward one side. Within the limits of these process variations, the overall bias is reduced by the back-to-back structure.

System Noise Analysis

One of the principal goals of this investigation was minimization of sensor self-noise. Although the internal gas composition and pressure were not known, the overall frequency response (with resonance at about 2500 Hz and a Q of less than 10) suggests a significant level of gas damping. Since the total moving mass of the full-bridge structure is 5 milligrams, the expected noise floor[3] set by molecular-thermal agitation (Eq. 2.11) is about 0.25 micro-g per root hertz. Other sources of noise are drive-oscillator amplitude noise, driver amplifier noise, preamplifier noise, and multiplier noise.

The drive and detection components are shown in the Fig. 5.7. A 45 kHz sine wave is generated by the XR2206 function-generator chip. The OP284 dual-operational-amplifier feeds the bridge with 0° and 180° copies of that sine wave. An INA141 instrumentation amplifier with a voltage gain of 10 amplifies the bridge output. This is followed by a passive (RC) high-pass filter (16 kHz) and an amplifier with a voltage gain of two using one half of the dual LF412. The amplified and filtered bridge output is then multiplied by the drive signal in the AD633 analog multiplier. The output of the multiplier then passes through a two-stage low-pass filter (1.6 kHz) using the other half of the LF412.

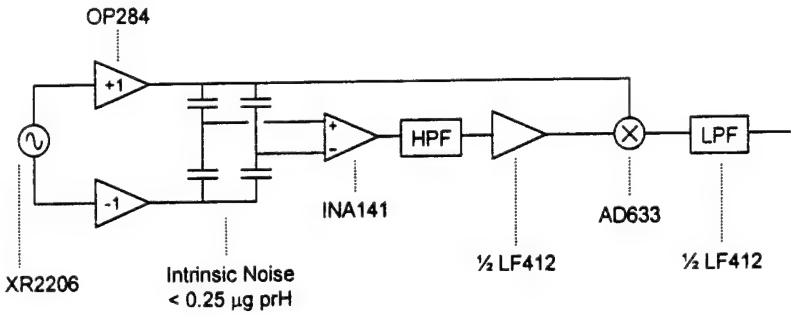

Figure 5.7. Major components of the drive and detection electronics used with the full-bridge, differential-capacitance accelerometer.

This is considerably more complex than that required by an ordinary passive sensor. Consequently, it is of considerable interest to examine the contributions of each component to the noise budget. For this purpose, we will refer all noise values to the sense-element output (i.e., the input of the INA141).

The XR2206 is an analog function generator. While the sine-wave output is not noise-free, the fluctuations are small and they are cancelled, to first order, by the bridge. These source fluctuations make only a negligible contribution to the overall noise. The drive amplifier pair (OP284) has an input voltage noise and an input current noise. The current noise is negligible because the impedance of the XR2206 output is low. The voltage noise is 4 nV per root hertz (prH). Since the magnitude of either of the drive-amplifier gains is unity, the effective noise translated to the bridge consists of two uncorrelated components of 4 nV prH or about 6 nV prH total.

The INA141 (set to x10 gain) has an input voltage noise of 12 nV prH and an input current noise of 0.3 pA prH. The current noise flows into the bridge, which has a terminal capacitance of 35 pF. At 45 kHz, the impedance is 100 KΩ, so the current noise contributes an equivalent voltage-noise component of 30 nV prH.

The Johnson noise of the passive RC high-pass filter is negligible compared to the input noise of the next gain stage (half of the LF412). The current noise of the LF412 is negligible and the voltage noise of 25 nV prH translates back to the input of the INA141 by dividing by the instrumentation amplifier's gain. The result is a contribution of 2.5 nV prH.

The noise of the analog multiplier[4] (AD633) is specified as an output noise spectral density of 800 nV prH. The system gain from sense element to the output of the multiplier is a voltage gain of 11, so the noise referred to the sense-element output is 73 nV prH. The low-pass filter after the multiplier has negligible contribution.

The acceleration response of this system is about 3 volts per g (see Fig. 5.11 in the next section). Consequently, the intrinsic noise of the MEMS structure translates to 750 nV prH at the output or (dividing by 11) 68 nV prH at the input to the instrumentation amplifier. The noise budget is summarized in Table 6.1.

Device	Voltage Noise nV prH	NEA micro-g prH
MEMS chip	68	0.25
OP284	6	0.02
INA141 (e_n)	12	0.04
INA141 (i_nZ_s)	30	0.11
LF412	2.5	0.01
AD633	73	0.27

Table 6.1. Noise budget for differential-capacitance accelerometer and drive circuitry. The noise components are shown both in equivalent voltage noise at the transducer and as noise-equivalent-acceleration (NEA). The root-mean-square sum yields a predicted NEA of 0.4 micro-g per root hertz (prH).

This system is close to reaching the intrinsic noise of the sense element. The dominant electrical noise component is that of the multiplier. If more gain were used prior to the multiplier, then its effective contribution could be reduced. At some point, the current-noise contributed by the INA141 would take over as the dominant component. This, then, could be reduced by increasing the drive frequency. However, for the microstructures used here, the zero-signal offset is too large (even with the full-bridge partial cancellation) to permit much more gain before the multiplier. This underscores the importance of repeatability in production of the microstructures. In the full-bridge implementation, it is not critical that the individual chips be accurately centered but it is important that the single-chip offset be consistent.

In order to reach the desired noise floor (between 10 and 100 nano-g prH from 5 to 1000 Hz), it would be necessary to modify both the electronics and the mechanical structure. Since the MEMS chip used here is far smaller than required for the ultimate package size, the most direct route to lower the noise is to make the structure larger to increase the mass of the moving element. If the moving-element mass is increased by a factor of 100 (7 mm by 7 mm by 2 mm in silicon; smaller in a more-dense metal), the physical size would still be acceptable but the noise floor would be reduced in amplitude by a factor of 10 (to 25 nano-g prH). In fact, this size is large enough to consider conventional machining so that a more dense material could be used for the moving element and a symmetric structure could be designed to reduce the zero-signal offset.

Evaluation

The bare MEMS chip is shown in Fig. 5.8. For each accelerometer element, two chips were bonded back-to-back and attached to a small printed-circuit board with the first-stage preamplifier in its small-outline (SOIC) version. The circuit board was mounted on a short acrylic stub as shown in the figure. This stub was used to mount the sensor in the calibration and noise measurement fixture. The stub is mounted in line with the sensor element but off center with respect to the printed circuit board assembly. This produced some minor resonances associated with circuit-board vibration. These resonances would not be present in the final package since the entire circuit board will be supported by an outer shell.

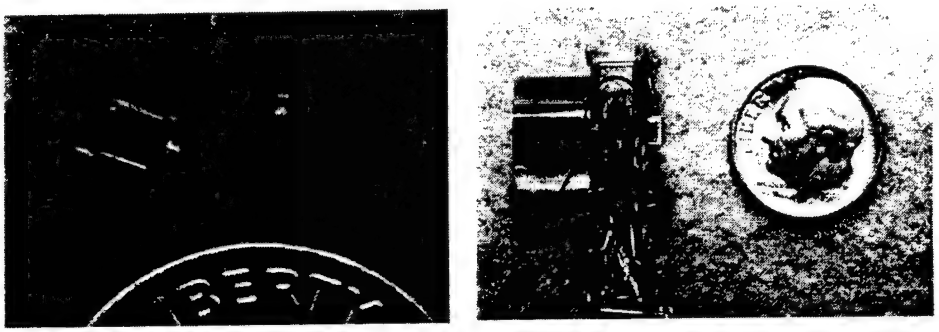

Figure 5.8. The bare MEMS chips are shown on the left. The bonded pair of chips is shown on the right mounted on the first-stage preamplifier board. The short acrylic stub is a temporary fixture for clamping the sensor in the calibration apparatus.

For the purposes of evaluation, the drive and detection electronics (except for the INA141) were built on a separate circuit board (Fig. 5.9).

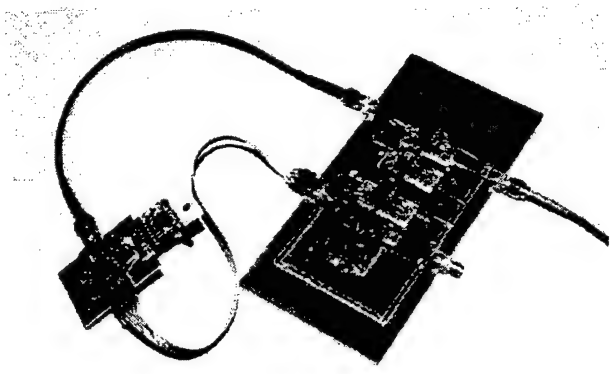

Figure 5.9. The drive and detection electronics are on the large board. An early version of the first-stage preamplifier and sensor head is shown to the left. This is replaced by the smaller package shown in Fig. 5.8.

The modular form was useful for investigating changes in the electronics. This also permitted long wires between the sensor element and the detection circuitry so that

electromagnetic interference could be intentionally introduced. The circuitry shown includes an earlier version of the first-stage preamplifier (the smaller board on the left).

The test sensor was calibrated on a pendulum by comparison with a reference accelerometer (Fig. 5.10). This apparatus was also used to produce an absolute calibration of the reference accelerometer by reciprocity[5].

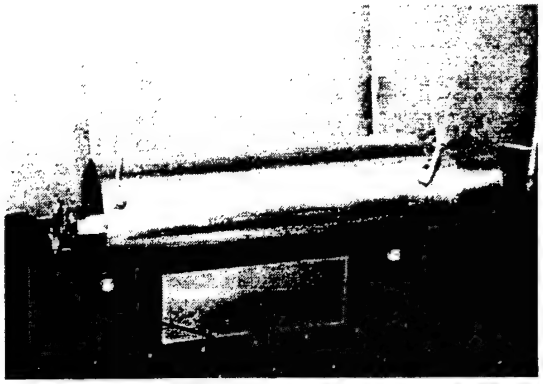

Figure 5.10. Pendulum apparatus for calibration and noise isolation. The differential-capacitance accelerometer is shown attached to the left end of the pendulum mass. A geophone is attached to the right end to drive the pendulum.

Measurement of the response of the sensor to acceleration is shown in Fig. 5.11. In both the magnitude and phase response, there is evidence of two small resonances between 800 and 1000 Hz. These are associated with flexure of the portion of the circuit board that extends out beyond the acrylic stub. The full hydrophone package will prevent such flexure. Disregarding these spurious resonances, the magnitude response is flat within one dB and the phase is flat within a few degrees to 1000 Hz.

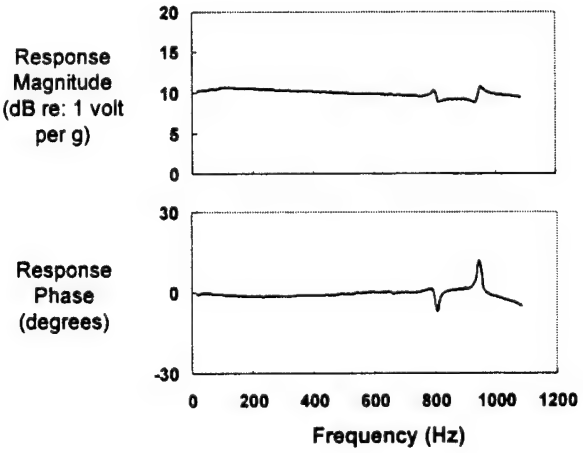

Figure 5.11. Magnitude and phase response of the differential-capacitance accelerometer. The small resonances between 800 and 1000 Hz are spurious resonances associated with circuit-board flexing.

The apparatus used for response calibration is a free-pendulum reciprocity calibrator, which also provides excellent mechanical isolation against external vibration. Consequently, the self-noise measurements were made on the apparatus immediately after the response calibration. A reference piezoelectric accelerometer was also used in the noise measurements.

As shown in Fig. 5.12, the full-bridge sensor has a self-noise level well below 1 micro-g per root hertz over all but the lowest part of the band. Normally, a high-frequency modulated sensor system would not show the 1/f behavior seen below 500 hertz here. However, as mentioned previously, the noise floor is dominated by the analog multiplier and its baseband 1/f noise is apparent. Notice the 60-Hz line in the upper trace and its absence in the full-bridge measurement. Although the reference accelerometer is well shielded and the prototype full-bridge sensor is not, the immunity to interference is much greater in the AC-drive sensor.

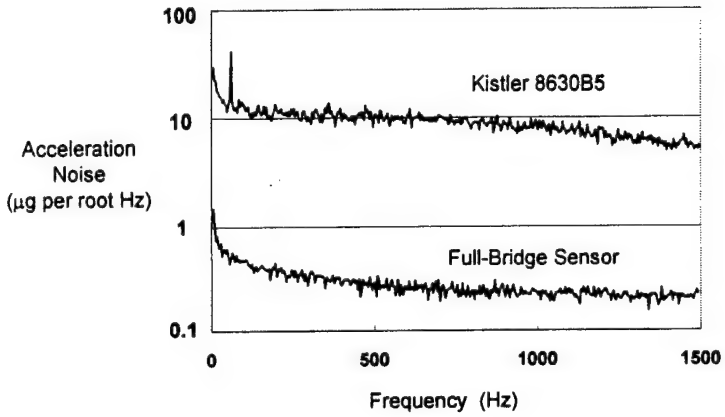

Figure 5.12. Equivalent acceleration noise for the full-bridge sensor. For reference, a piezoelectric accelerometer of similar packaged volume is also shown.

The measured noise floor (0.3 micro-g per root hertz or less at the upper end of the frequency band in Fig. 5.12) is slightly below the predicted value (0.4 micro-g per root hertz from Table 6.1). Manufacturer's specifications for "typical" noise performance were used rather than the "minimum" values so it is likely that the multiplier chip was somewhat below its "typical" specification. The estimate for the intrinsic noise of the MEMS chip is crude and may be an overestimate. The chip intrinsic noise and the multiplier noise dominate the noise budget; errors in the other components would have little effect.

The full-bridge structure also permits design of multiple-axis accelerometers with a common phase-center. The two-axis version of the structure is assembled as in Fig. 5.13. The upper and lower chips form the full bridge for sensing along one axis while the left and right chips form the full bridge for sensing along the perpendicular axis.

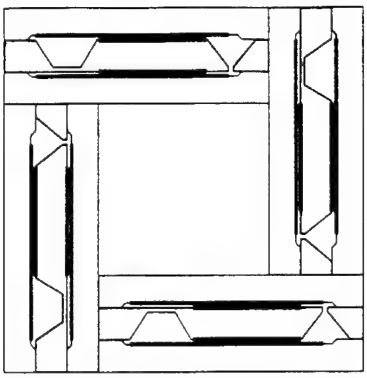

Figure 5.13. Structure for two-axis, full-bridge accelerometer. The chips are connected in opposing pairs. The effective phase center is identical for both axes.

The hydrophone element does not require an extremely large dynamic range but it is useful to examine the limits to dynamic range for this sensor. One inherent limitation is that the differential capacitor is not linear. As the center mass approaches either side plate, the effects of electrostatic field fringing introduce substantial nonlinearity. This can, in principle, be reduced through the use of guard-ring electrodes.

The dynamic range of the AD633 analog multiplier is specified by the manufacturer to be 140 dB and can be used up to 1 MHz. (Actually, it can be used as a demodulator well above 1 MHz since only the baseband output of the multiplication is important.) Switching demodulators are available for higher frequency carriers and can achieve dynamic ranges in excess of 100 dB with simpler, square-wave drive.

The multiplication can be done digitally to minimize the noise introduced at this stage. The dynamic range is then limited by the word size and the multiplication algorithm. The differential-capacitance sensor is amenable to direct sigma-delta conversion but the implications for noise floor have not been examined.

Multiplexing AC-Drive Sensors

Commercial differential-capacitance accelerometers combine the drive oscillator and the detector in the same package as the sense element. This makes the internal architecture transparent to the user but prevents exploitation of some of the potential (beyond low noise) of AC-biased sensors. As discussed above, if the detection is deferred, then substantial immunity to external electromagnetic interference is obtained. In this case, it is more convenient to supply the AC-drive signal externally also (since it must be supplied to the detection process). If the sensor package consists of a divide-by-N module with positive- and negative-phase output and the sense element (see Fig. 5.14), then several sensors can be multiplexed while, at the same time, preserving the external-interference immunity.

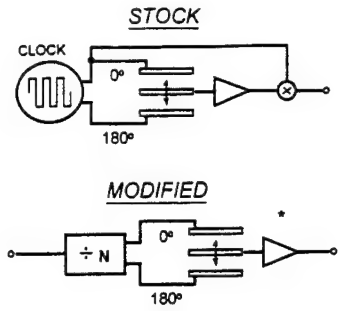

Figure 5.14. The standard architecture includes the drive and detection electronics in the same package with the sensor (top). The architecture can be modified for multiplexing by supplying an external clock and detecting after demultiplexing. The divide-by-N module provides unique coding for each sensor.

Fig. 5.15 illustrates one possible multiplexing architecture. The outputs from three sensors, each having a different divide-by-N module, are directly summed with the clock signal.

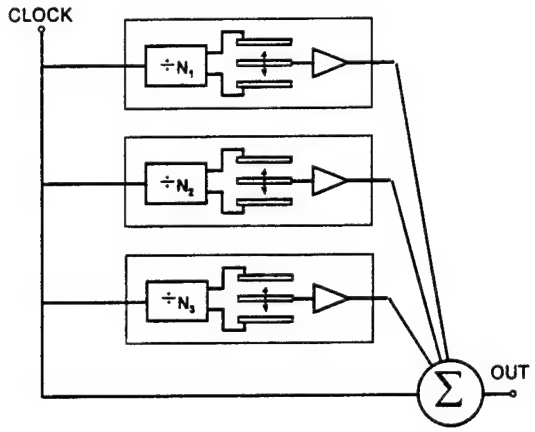

Figure 5.15. Architecture for multiplexing. The master clock and three modulated signals are summed onto the transmission line. Each output is separable at the receiving end by reversing the process.

The master clock can be extracted by narrow-band filtering and a phase-locked loop at the "receiver" end of the system. Once the clock is recovered, the divide-by-N process is repeated and synchronous demodulation extracts the individual sensor outputs.
Thus, the AC drive provides a natural means for implementing frequency-division multiplexing. This advantage is preserved even if the signals are transmitted optically.

[1] This assumes that the "carrier" of the sense signal is in phase with the drive frequency. In this implementation, there is a small phase difference between the carrier and the reference signal. The effect is to decrease the magnitude of the detected signal but only by the cosine of that phase difference. The phase difference can be

removed but the marginal gain in performance is not worth the complication. For much higher frequency drive, the phase difference might be large enough to warrant correction.

[2] CSEM Microsystems, Centre Suisse d'Electronique et de Microtechnique SA, P.O. Box CH-2007, Neuchâtel, Switzerland. Various commercial products are mentioned in this report but none of these references should be construed as endorsements or recommendations.

[3] Bear in mind that the SI unit for acceleration is m/s^2. The more conventional unit of "g" is used here. One g is 9.8 m/s^2.

[4] At and beyond the multiplier, the relevant noise is the noise in the band of desired signals. Prior to the multiplier, the relevant noise is the noise at 45 kHz.

[5] T. Gabrielson, "Free-mass reciprocity calibration," J. Acoust. Soc. Am. **102**, 2800-2808, 1997.

Chapter 6

Velocity Sensor Suspension Dynamics

Velocity sensors differ in two critical aspects from pressure hydrophones. Velocity sensors are inherently directional and a velocity sensor must be sufficiently free to respond faithfully to the local acoustic particle velocity. The latter aspect is the subject of this chapter. A pressure hydrophone can be rigidly fixed in place and, as long as the effects of diffraction are properly accounted, the pressure response is an accurate representation of the local acoustic pressure. If a velocity sensor were rigidly fixed in space, the output would be zero regardless of the local acoustic field.

The following conditions must be true for accurate velocity sensing:

- The sensor and any mount or suspension must not perturb the acoustic field
- The sensor must respond with good fidelity to the local particle velocity
- The sensor must be isolated from vibration transmitted through the suspension
- The sensor must be protected from non-acoustic fluid motion (turbulence, for example)

Velocity sensors are considerably more sensitive to motional disturbance than pressure sensors. Even those pressure hydrophones designed with no regard for acceleration canceling, do not respond strongly to motion. The velocity sensor is expressly designed to sense motion and it cannot discriminate between vibration transmitted through its mount or flow turbulence and acoustic particle velocity. Isolation that is good enough for pressure hydrophones may be entirely inadequate for velocity sensors.

In order to consider the practical issues associated with the effective use of velocity sensors, we will develop the fundamental dynamics of the suspended velocity sensor. While there are systems for which suspension is not an issue (accelerometers in neutrally buoyant, free-floating housings, for example), these applications are exceptional. In most applications, an inadequate suspension can render the best velocity sensor useless. Suspension design should be a fundamental and early component in the development of a velocity sensor system.

Basic Velocity-Sensor Dynamics

In its most basic form, the acoustic velocity sensor is an inertial sensor structure inside a housing that is free to react to the acoustic velocity field. Ideally, the sensor housing would be much less than an acoustic wavelength in dimension, completely rigid[1], and neutrally buoyant. In Fig. 6.1, the sensor is represented as a damped mass-spring oscillator. The incident acoustic field produces a force on the housing or case; the case responds to the acoustic field; and the sensor produces an output related to the relative displacement between the proof mass and the case. The case "sees" an impedance with a reactive component primarily associated with induced water mass – that is, water that is entrained by the case and that moves with the case.

The real part of the impedance has components related to the hydrodynamic drag of the case and to acoustic radiation from the case.

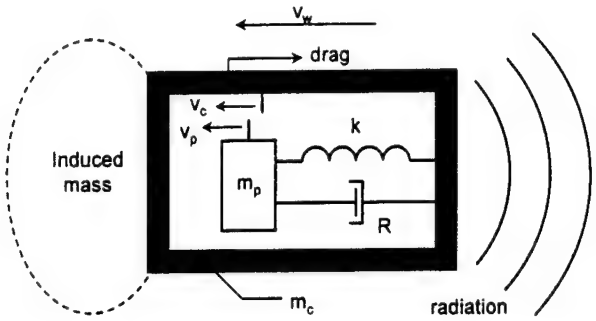

Figure 6.1. A velocity sensor immersed in a fluid. The dynamics of the inertial sensing structure are represented by a damped mass-spring system. The motion of the housing is influenced by the incident acoustic field, hydrodynamic forces (induced mass and drag), and acoustic re-radiation.

These actions and reactions can be represented by an electrical equivalent circuit[2] (see Fig. 6.2). The force, F, is the integral of the acoustic pressure over the case, the acoustic particle velocity (the component aligned with the sense axis of the sensor) is v_w, and the velocity of the case is v_c. The masses are represented as inductors. The proof mass moves with velocity, v_p. The spring and the mechanical damper react to the difference in displacement or velocity between the case and the proof mass.

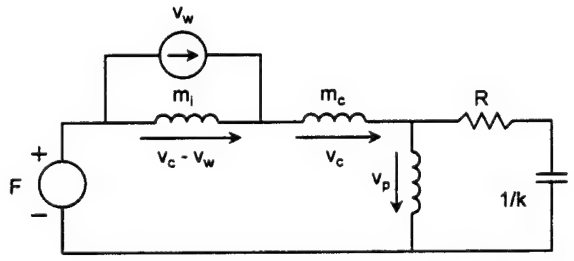

Figure 6.2. Equivalent circuit for immersed sensor body including the hydrodynamic induced mass, m_i. Mass elements are represented by inductors, stiffness by a capacitor, and mechanical damping by a resistor.

The circuit representation in Fig. 6.2 ignores hydrodynamic drag on the body but does include the more important hydrodynamic effect of induced mass. The influence of the induced mass depends on the difference in velocity between the case and the water, hence the form in Fig. 6.2. If the case were moving in perfect concert with the water ($v_c - v_w = 0$), then there would be no inertial force associated with entrained water mass.

The magnitude of the induced mass can be determined from the radiation impedance. For simplicity, we will consider a spherical body with radius, a. The radiation impedance is[3]

$$Z_{rad} = \rho_w c_w \pi a^2 \left[\frac{(ka)^4}{3} + j\frac{2ka}{3} \right] \tag{6.1}$$

where ρ_w and c_w are the density and sound speed of water and k is the wave number ($= \omega/c_w$). The reactance is mass-like (having the form $+j\omega$ times a mass) where the mass is the induced mass, m_i,

$$m_i = \rho_w 2\pi a^3/3 = \rho_w V_c/2 \tag{6.2}$$

Here, V_c is the volume of the spherical case. It is convenient to express the mass of the case in terms of its equivalent density, ρ_c, and volume:

$$m_c = \rho_c 4\pi a^3/3 = \rho_c V_c \tag{6.3}$$

(The "case" includes all internal parts except the proof mass, which is assumed small.) The force exerted on the case by the acoustic field can be determined by integrating the pressure field over the surface of the sphere; however, for $ka << 1$, it is approximately equal to the pressure gradient times the volume. The pressure gradient, which is force per unit volume, is

$$-\frac{\partial p}{\partial x} = \rho_w \frac{\partial v_w}{\partial t} = j\omega\rho_w v_w \tag{6.4}$$

Assuming that the particle-velocity vector is aligned with the sense axis, the force on the case is

$$F \approx -\frac{\partial p}{\partial x} V_c = j\omega\rho_w V_c v_w \tag{6.5}$$

Ignoring the internal dynamics of the sensor element, the solution for the ratio between the case velocity and the acoustic velocity is[4]

$$\frac{v_c}{v_w} = \frac{3\rho_w}{\rho_w + 2\rho_c} \tag{6.6}$$

If the aggregate density of the case (i.e., total mass divided by total volume) is equal to the density of the water, then the case velocity is identical to the acoustic particle velocity. If the case is infinitely massive, it does not move; if the case density is negligible compared to that of water, then the case moves three times as fast as the particle velocity. In practice, velocity sensors may be intentionally designed to have densities different from that of water. In those cases, the appropriate correction must be made to the velocity response.

If we consider the full dynamics of the sensor element (R, k, and m_p), then the ratio of the output quantity (the relative velocity between case and proof mass) and the particle velocity is

$$\frac{v_c - v_p}{v_w} = \frac{\alpha}{1 - \gamma\left(\Omega^2 + j\,\Omega/Q\right)} \tag{6.7}$$

where

$$\alpha = \frac{m_d + m_i}{m_c + m_i} \tag{6.8}$$

$$\gamma = \frac{m_p + m_c + m_i}{m_c + m_i} \tag{6.9}$$

$$\Omega = \frac{\omega_0}{\omega} \tag{6.10}$$

and

$$Q = \frac{\omega_0 m_p}{R} \tag{6.11}$$

In the definitions above, m_d is the mass of water displaced by the case, and ω_0 is the resonance frequency of the sensor spring-mass system.

If the actual transduction mechanism produces an output directly proportional to the relative displacement between the proof mass and the case (as a capacitive transducer might), then, below the fundamental resonance, that output will be directly proportional to the case acceleration. We will consider only accelerometer-based sensors in what follows; however, the treatment is easily adapted to other sensor types. The ratio of sensor output to acoustic particle acceleration is then

$$\frac{x_c - x_p}{a_w} = -\frac{1}{\omega^2}\frac{v_c - v_p}{v_w} = -\frac{\Omega^2}{\omega_0^2}\frac{v_c - v_p}{v_w} \tag{6.12}$$

For frequencies well below the sensor resonance (large Ω), the magnitude of the response ratio from Eqs. 6.7 and 6.12 is

$$\left|\frac{x_c - x_p}{a_w}\right| \approx \frac{\alpha}{\gamma\,\omega_0^2} \tag{6.13}$$

and the phase is

$$\phi \approx -\frac{1}{\Omega Q} \tag{6.14}$$

(The acoustic particle velocity can be obtained directly from the particle acceleration by dividing the acceleration by ω.) For a critically damped accelerometer ($Q = 0.5$), a decade below the resonance ($\Omega = 10$), the phase error would be about 11 degrees. At a given frequency, the phase error decreases with increasing Q. The magnitude of this phase error is an important consideration in velocity sensor design. The phase error can be reduced by increasing the Q, or by moving the resonance frequency further from the band of interest. The error can also be compensated if the sensor is individually calibrated for phase response.

If an accelerometer is housed in a rigid case and the entire unit is neutrally buoyant, then (as long as $ka << 1$) the output will accurately reflect the local particle acceleration. In hydrophone design, coatings and housings are often made from polymers that have a ρc product close to that of water. This is intended to provide an accurate free-field pressure at the sense-element face. *This design philosophy is entirely irrelevant for the velocity sensor* as its fundamental response is inertial.

Influence of the Suspension

Except for truly free sensors[5], a velocity sensor normally needs to be held in a fixed average position in the acoustic field. Many of the practical problems in velocity-sensor implementation stem from this requirement. The finest velocity sensor element can be rendered ineffective by poor suspension.

We can modify the equivalent circuit of Fig. 6.2 to include the suspension as shown in Fig. 6.3. Here, the sensor-element dynamics have been lumped into an impedance element, Z_x. Two additional elements representing the suspension stiffness, k_s, and the suspension damping, R_s, are placed in a branch that "sees" the velocity difference between the case and the suspension "fixed" point. (The suspension fixed point is permitted to move with velocity, v_s, in order to analyze the transmission of disturbances from the support structure through the suspension.)

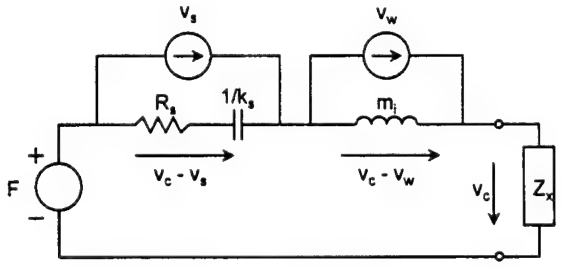

Figure 6.3. Equivalent circuit including the suspension dynamics.

Assume that the sensor is neutrally buoyant and the frequency of interest is well below the sensor resonance. There are two points of interest: one is the fidelity with which the suspended sensor will reproduce the acoustic signal; the other is the sensitivity of the sensor to motion in the suspension mount.

For the first case, set v_s to zero. Then the ratio of the case velocity to the water velocity is

$$\frac{v_c}{v_w} = \frac{1}{1 - \frac{k_s}{\omega^2 m_c} + \frac{R_s}{j\omega m_c}} \tag{6.15}$$

Ideally, this ratio should be one. The ratio, k_s/m_c, is the square of 2π times the resonance frequency of the suspension/case system. In order for the case to respond accurately to the acoustic field, this resonance must be well below the frequency of interest. Both the magnitude and phase are important. Since a suspension normally has a rather large damping, the phase error will still be significant well above the resonance. This can be problematic for low-frequency sensors.

To examine isolation from the suspension mounting point, set F to zero. The ratio of case velocity to suspension mount-point velocity is then

$$\frac{v_c}{v_s} = \frac{1}{1 - \frac{\omega^2 m_c / k_s}{1 + j\omega R_s / k_s}} \tag{6.16}$$

This quantity must be small enough so that the expected levels of structural vibration are reduced to acceptable levels. The criterion is the same as in the first case. For this quantity to be small, the resonance frequency of the suspension/sensor system must be much lower than the signal frequency of interest.

Types of Suspensions

In designing suspensions, the specific type of suspension will determine how the stiffness is calculated. There are three primary varieties of velocity-sensor suspensions. These are shown diagrammatically in Fig. 6.4.

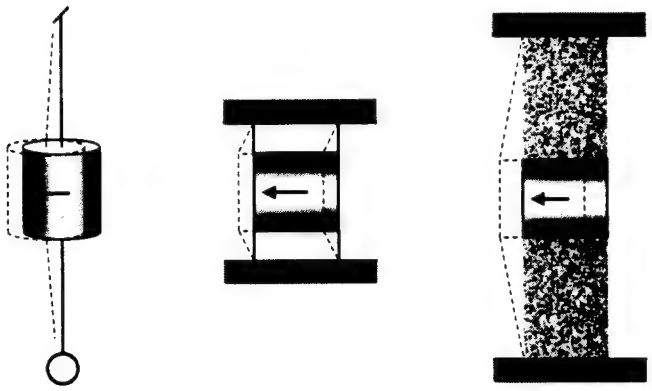

Figure 6.4. Primary varieties of velocity-sensor suspension. The suspended (sonobuoy) sensor is shown on the left. A sensor supported by mechanical springs at each end is shown in the center and a sensor embedded in a compliant suspension material is shown on the right.

On the far left is the standard suspension used for the two-axis velocity sensor in the directional (DIFAR) sonobuoy. The two axes of sensing are in the horizontal plane. The suspension consists of the signal cable above and a weight below. In some versions, the lower weight is not used; instead, the housing is negatively buoyant. The restoring force of this suspension in the horizontal direction is very small for small motion. Weak restoring force in the sense direction is critical for a low suspension resonance. The restoring force in the vertical direction is much higher but no sensing is done in the vertical direction.

In the center of Fig. 6.4, a single-axis sensor is shown supported by mechanical springs at each end. Circular, folded-cantilever springs (as used internally in geophones) provide weak force axially and stronger force laterally. This is a very effective, compact suspension but it is more difficult to extend to two-axis systems.

A third option is shown on the far right in Fig. 6.4. Here, the sensor is embedded in a compliant material. If the compliant material extends well beyond the sensor before its attachment to rigid structure, then the material could be analyzed as a second acoustic medium; however, this is rarely practical. Frequently, the material's extent from sensor to support is less than a wavelength for shear waves in the material, so it acts like a restoring spring for displacement of the sensor case.

For an isotropic, elastic material, the shear modulus, G, is related to the Young's modulus, E, by the Poisson's ratio, σ:

$$G = \frac{E}{2(1 + \sigma)} \tag{6.17}$$

The bulk modulus, B, is given by

$$B = \frac{E}{3(1 - 2\sigma)} \tag{6.18}$$

and the ratio of shear modulus to bulk modulus, which is also the square of the ratio of the shear sound speed to the compressional sound speed, is

$$\frac{G}{B} = \frac{c_{shear}^2}{c_{compr}^2} = \frac{3(1 - 2\sigma)}{2(1 + \sigma)} \tag{6.19}$$

The shear speed is zero for an ideally incompressible ($\sigma = 0.5$) material but we are interested in materials that have sound speeds within a factor of two of that for water[6]. Such materials may have Poisson's ratios as low as 0.48 corresponding to a G/B ratio of 0.04. The shear wave speed, then, would be about one-fifth of the compressional speed in the material.

The shear stiffness of the material is only negligible if the shear wavelength is much less than the distance from the sensor body to the support. From the analysis above, the shear wavelength may be one-fifth of the compressional wavelength. At the high end of the frequency range, the sensor body might be one-tenth the compressional wavelength or one-half the shear wavelength long. If the compliant material extends to ten times the sensor dimension, then, at the upper range of frequency, the suspension would be a few shear wavelengths in extent. However, at lower frequencies, the shear restoring force would need to be considered. In addition, the shear loss in such materials is normally high, which further affects the response of the embedded sensor body.

Suspension Anisotropy

At this point, the design problem of velocity-sensor suspension might be considered fundamentally solved. However, we have only treated the problem in one dimension. If the suspension has a stiffness that depends strongly on the direction of force application, we must also consider the direction of motion of the sensor body relative to the direction of the velocity field. If the suspension has directional anisotropy, then the sensor body will not move in the same direction as the direction of the particle velocity vector.

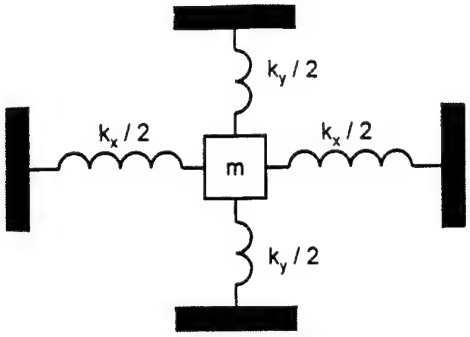

Figure 6.5. A two-dimensional suspension. The spring constants in the two orthogonal directions have different stiffnesses.

Consider the schematic suspension shown in Fig. 6.5. The stiffness in the x-direction is k_x and the stiffness in the y-direction is k_y. The components of the displacement, D, in the x- and y-directions are related to the x- and y-components of the applied force, F, as follows:

$$D_x = k_x^{-1} \cdot F_x + 0 \cdot F_y \tag{6.20}$$

$$D_y = 0 \cdot F_x + k_y^{-1} \cdot F_y \tag{6.21}$$

If k_x is not equal to k_y, it is clear that the displacement is in a different direction than the applied force.

In matrix form, the vector displacement, **D**, is related to the vector force, **F**, by the compliance matrix, **c**:

$$\mathbf{D} = \mathbf{c} \cdot \mathbf{F} \tag{6.22}$$

where

$$\mathbf{c} = \begin{bmatrix} k_x^{-1} & 0 \\ 0 & k_y^{-1} \end{bmatrix} \tag{6.23}$$

The matrix rotation operator, **A**, for two dimensions is

$$\mathbf{A} \equiv \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \tag{6.24}$$

This operator rotates the coordinates by the angle, ϕ. We can operate on Eq. 6.22 to examine the effects on the compliance matrix. Multiply both sides by the rotation operator,

$$\mathbf{A} \cdot \mathbf{D} = \mathbf{A} \cdot \mathbf{c} \cdot \mathbf{F} = \mathbf{A} \cdot \mathbf{c} \cdot \mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{F} \tag{6.25}$$

and rewrite the equation as a new equation in the rotated coordinate system:

$$\mathbf{D'} = \mathbf{A}\cdot\mathbf{c}\cdot\mathbf{A}^{-1}\cdot\mathbf{F'} = \mathbf{c'}\cdot\mathbf{F'} \tag{6.26}$$

where the rotated compliance matrix is

$$\mathbf{c'} = \begin{bmatrix} k_x^{-1}\cos^2\phi + k_y^{-1}\sin^2\phi & \left(k_y^{-1} - k_x^{-1}\right)\sin\phi\cos\phi \\ \left(k_y^{-1} - k_x^{-1}\right)\sin\phi\cos\phi & k_y^{-1}\cos^2\phi + k_x^{-1}\sin^2\phi \end{bmatrix} \tag{6.27}$$

The displacement will only be in the same direction as the force if the force is applied along the direction of a coordinate axis for which the compliance matrix is diagonal. An arbitrary force only generates a displacement parallel to that force if the compliance matrix is diagonal for any rotation angle. From Eq. 6.27, this is only true if $k_x = k_y$. If the two stiffnesses are not identical, then the situation shown in Fig. 6.6 arises in which a force applied in one direction produces a displacement in another direction. For a velocity sensor, this would cause the apparent direction (and magnitude) of the velocity to be in error.

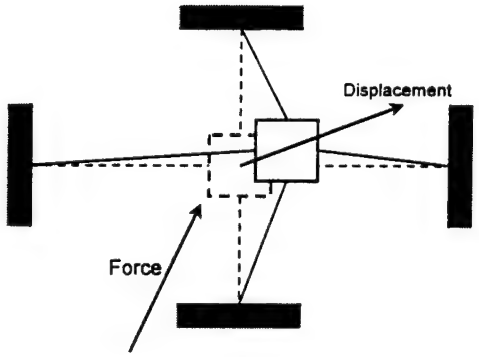

Figure 6.6. Effects of suspension anisotropy. The displacement caused by an applied force is not in the same direction as the force but tends to be rotated toward the weaker axis.

Once more, consider the representative suspensions shown in Fig. 6.4. The sonobuoy suspension has virtually identical stiffness in the two horizontal coordinate directions so the direction of the projection of the particle velocity vector onto the horizontal plane will be determined accurately. The stiffness in the vertical direction is substantially greater then the stiffness in the horizontal direction so a vector force inclined to the horizontal plane will produce a displacement principally in the horizontal plane – the vector will appear to have been pushed down toward horizontal. Since the sonobuoy is intended to determine horizontal bearing, this is of little consequence. The most pronounced effect is to increase the level of the apparent signal for off-horizontal arrivals, which is beneficial as long as the vertical angle of arrival is irrelevant.

For either of the other two suspensions shown in Fig. 6.4, it is straightforward to arrange the stiffnesses in both directions *perpendicular* to the sense axis to be equal but it is much more difficult to arrange the on-axis stiffness to be equal to the perpendicular stiffness. Consequently,

off-axis arrivals will have produce errors in the magnitude of the velocity. If the on-axis stiffness is less than the off-axis stiffness, then the vector will appear to be rotated toward the sense axis and the magnitude will be overestimated.

If a vector sensor is constructed from two or more single-axis sensors of this type, the reconstructed direction will generally be in error. If the direction of arrival is parallel to either sensor or it exactly bisects the sensor axes, the measured direction will be accurate. In all other cases, the anisotropic stiffness biases the direction determination. While the direction is determined correctly for arrivals along the bisector, the magnitude is not because both sensors overestimate the magnitude.

Incidentally, these considerations suggest that the practice of measuring response only along the principal axes of the sensor is inadequate. Excitation along the principal axes will not show the errors that result from off-axis excitation.

Furthermore, there is a connection between design symmetry and performance. For the configurations discussed above, it is clear that if the sensor-plus-suspension has an axis of symmetry, then arranging for suspension isotropy in the plane perpendicular to the symmetry axis is straightforward. Because of this, it is considerably easier in practice to maintain high quality in a two-axis design than in a three-axis design.

Impedance-Matching Material

Before considering arrays of sensors embedded in a compliant material, it is instructive to review the properties of impedance-matching ("ρc") materials. As mentioned previously, there is no benefit to embedding velocity sensors in such material from the standpoint of enhancing their response. However, if the density of such material is less than that of water, the ρc-material can be used to compensate for excess density in the sensor body to achieve neutral buoyancy. Alternatively, the material can be used as a compliant suspension.

Impedance-matching material is often considered transparent to an acoustic field. This is only true in the special case in which the incident acoustic wave is normal to the material and the material is sufficiently free that edge constraints are inconsequential. Even if the material can be treated as a fluid, the reflection coefficient depends on the incident angle, θ_i, as follows:

$$R = \frac{\rho_2 c_2 \cos\theta_i - \rho_1 c_1 \sqrt{1 - (c_2/c_1)^2 \sin^2\theta_i}}{\rho_2 c_2 \cos\theta_i + \rho_1 c_1 \sqrt{1 - (c_2/c_1)^2 \sin^2\theta_i}} \tag{6.28}$$

If the incident angle is zero and $\rho_1 c_1 = \rho_2 c_2$, the reflection coefficient is zero – the energy passes from fluid 1 (water, in our case) entirely into fluid 2 (the compliant material). If water is also on the other side of a planar layer of fluid 2, the energy also passes, without internal reflection, into the water on the far side. If the acoustic wave is not normally incident, however, then there is partial reflection and partial transmission.

Furthermore, the angle changes from fluid 1 to fluid 2 according to Snell's Law:

$$\frac{\sin\theta_1}{c_1} = \frac{\sin\theta_2}{c_2} \tag{6.29}$$

Consider a sphere of ρc-material in water. If the ρc-product is matched but the sound speed is different, then the density will also be different. Normally, the density is lower than water and the sound speed is higher. In this case, the ρc-sphere would be positively buoyant and move more than the local acoustic particle velocity (see Eq. 6.6). This could be used to compensate for excessive suspension constraint but the compensation may be frequency dependent and difficult to reproduce in practice.

Another aspect of ρc-material is the acoustic scattering from a volume of that material. The assumption that, since it is impedance matched to water, it does not scatter is incorrect. If an object's density differs from that of water (regardless of its compressibility), then it responds to an acoustic field with a velocity that differs from the local particle velocity. The relative difference in velocity causes scattering. Furthermore, if the compressibility differs, then the object's expansion and contraction do not match the corresponding expansion and contraction of the surrounding fluid and this, too, causes scattering. (Much of the success of common ρc–material results from the fact that neither the sound speed nor the density differs very much from those quantities in water.)

This entire discussion assumes that the ρc-material behaves as a fluid. In fact, such material has shear stiffness and this must be accounted in determining the interaction between an embedded velocity sensor, the material, and the mounting points. Beyond this, it may also be necessary to consider vibrational modes in the ρc-material itself.

Arrays of Velocity Sensors in a Compliant Layer

If a compliant suspension is acceptable (given the anisotropic response to velocity), then it is also possible to create an array of sensors by embedding individual sensors in a layer of compliant material. For such designs, the sensor-to-sensor interaction must be considered. The aggregate effects of periodic arrays of inclusions in a matrix are well known[7].

While the full theory for a two-dimensional lattice is complicated, the fundamental principles can be seen with a one-dimensional model. Fig. 6.7 shows a one-dimensional slice through a compliant matrix with embedded velocity sensors. This can be modeled as a periodic system of masses and springs where m_a represents the sensor mass, m_b represents the effective (dynamic) mass of the matrix, and k represents the effective stiffness of the matrix.

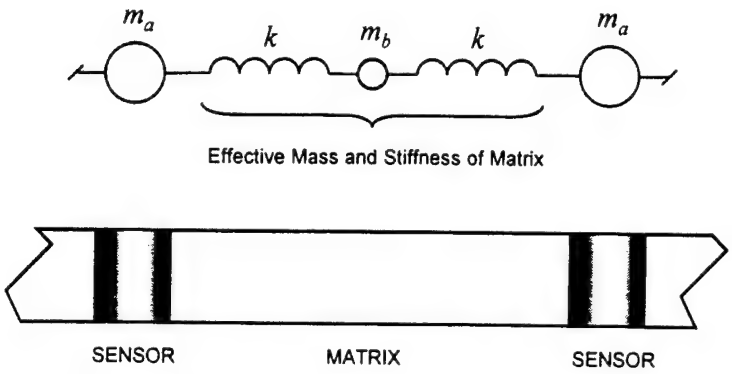

Figure 6.7. Cross-section of an array of sensors in a compliant matrix. A simple model for the behavior of such an array is shown above. The sensors are modeled as masses and the intervening matrix is modeled as an effective point mass and stiffness elements.

The relative motion of the sensor bodies with respect to the matrix can be either transverse or in-plane; in either case, the argument is qualitatively the same. For this one-dimensional, two-mass lattice, there are two distinct bands of conduction of waves through the lattice and two distinct "stop" bands in which waves decay exponentially without propagation.

At very low frequency, the sensor bodies and the matrix move in phase and both the sensor mass and the matrix center-of-mass move. As the frequency of excitation is increased, eventually the sensor bodies move such that the centers-of-mass of the matrix elements in between are stationary (assuming that the mass of the sensor body is greater than the effective mass of the intervening matrix element). Excitation at somewhat higher frequency only affects the elements in the immediate vicinity of the excitation; disturbances are attenuated exponentially into the lattice. As the frequency is increased still further, disturbances begin to propagate again but now the sensor bodies and the matrix elements move out-of-phase. At still higher frequency, disturbances localize once more and no further propagating waves are observed.

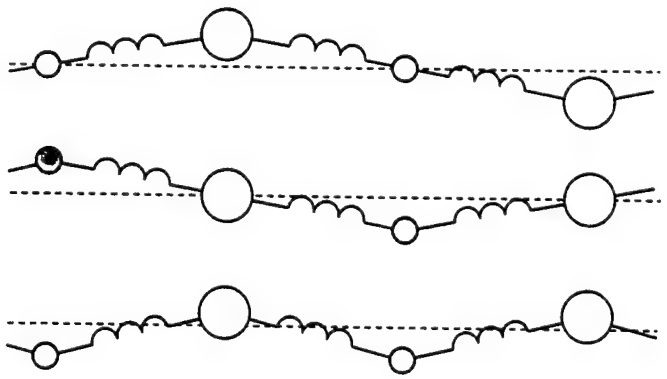

Figure 6.8. Representative motion of sensors and matrix near the edges of the two conduction bands. The uppermost diagram represents the motion at the upper frequency of the lower band. The lower two diagrams represent the motion at the lower and upper frequencies of the upper conduction band.

The bounding frequencies for these bands of conduction and exponential attenuation are given by the circumstances shown in Fig. 6.8. The lower conduction band extends from zero frequency to the resonance frequency of the sensor body when the centers-of-mass of the matrix elements between are stationary. The topmost picture in Fig. 6.8 shows the motion at the upper end of the lower conduction band where the large masses are moving alternately out-of-phase and the small masses (the matrix mass elements) are not moving. The lower conduction band covers the following frequency range:

$$0 < \omega^2 < \frac{2k}{m_a} \tag{6.30}$$

The upper conduction band starts at the frequency of resonance for the matrix center-of-mass motion if the sensor bodies are stationary. It extents to the frequency of resonance in which both masses move in opposition about their common center of mass. This band covers the frequency range:

$$\frac{2k}{m_b} < \omega^2 < \frac{2k(m_a + m_b)}{m_a m_b} \tag{6.31}$$

The two bounding frequencies for this band have the characteristic motion shown in the center and bottom diagrams in Fig. 6.8.

When designing an array of sensors in a compliant layer, it is important to design the distribution of elements according to the expected signal frequencies so that the expected signal excitation falls in either of the stop bands. Excitation in either conduction band will result in undesireable sensor-to-sensor interaction ("crosstalk").

These design principles are analogous to the design principles for 1-3 piezoelectric composite material[8]. For arrays built from 1-3 composite, if the elements must function individually, the material is designed so that the operating frequencies are in one of the stop bands. On the other hand, if the material is designed so that all of the individual elements act in concert, the material is designed so that the operating frequencies are in the lower conduction band. In that way, the matrix moves in phase with the ceramic elements and augments the motion.

---

[1] Designers of pressure hydrophones often include waterproof coatings that are nominally transparent to the acoustic pressure field. For a velocity sensor, the issue of acoustic transparency is irrelevant.

[2] T. Gabrielson, D. Gardner, and S. Garrett, "A simple neutrally buoyant sensor for direct measurement of particle velocity and intensity in water," J. Acoust. Soc. Am. 97, 2227-2237, 1995.

[3] P. Morse and U. Ingard, *Theoretical Acoustics*, Princeton University Press, Princeton, NJ, 1968.

[4] L. Landau and E. Lifshitz, *Fluid Mechanics*, Pergamon, NY, 1959, §11.

[5] For example: D. D'Spain, W. Hodgkiss, and G. Edmonds, "The simultaneous measurement of infrasonic acoustic particle velocity and acoustic pressure in the ocean by freely drifting Swallow Floats," IEEE J. Oceanic Eng. **16**, 195-207,1991; or the Neutrally Buoyant Infrasonic DIFAR Sonobuoy (NBIDS).

[6] True rubber has a Poisson's ratio of 0.4997. The harder polymers used for impedance matching are somewhat lower and acrylics are in the neighborhood of 0.40.

[7] L. Brillouin, Wave Propagation in Periodic Structures, McGraw-Hill, NY, 1946.

[8] W. Smith, A. Shaulov, and B. Auld, "Design of piezocomposites for ultrasonic transducers," Ferroelectrics **91**, 155-162, 1989.

Appendix A

Publications, Patents, and Presentations

PUBLICATIONS:

Gabrielson, T. B., "Frequency constants for transverse vibration of annular disks," J. Acoust. Soc. Am. **105**, 3311-3317, 1999.

Gabrielson, T. B., "Free-mass reciprocity calibration," J. Acoust. Soc. Am. **102**, 2800-2808, 1997.

Gabrielson, T. B., D. L. Gardner, and S. L. Garrett, "A simple neutrally buoyant sensor for direct measurement of particle velocity and intensity in water," J. Acoust. Soc. Am. **97**, 2227-2237, 1995.

Gabrielson, T. B., "Fundamental noise limits for miniature acoustic and vibration sensors," ASME J. Vibration and Acoustics **117**, 405-410, 1995.

Gabrielson, T. B., "Modeling and measuring self-noise in velocity and acceleration sensors," Proc. Acoustic Velocity Sensors Focused Workshop, Acoust. Soc. Am., 1995.

Rockstad, H., T. Kenny, J. Reynolds, W. Kaiser, and T. Gabrielson, "A miniature high-sensitivity broad-band accelerometer based on electron tunneling transducers," Sensors and Actuators A **43**, 107-114, 1994.

Gabrielson, T. B., "Mechanical-thermal noise in micromachined acoustic and vibration sensors," IEEE Trans. Electron Devices **ED-40**, 903-909, 1993.

Gabrielson, T. B., "Mechanical-thermal noise in microminiature hydrophones," in *Transducers for Sonics and Ultrasonics*, (ed. M. McCollum, B. Hamonic, and O. Wilson, Technomic Publ., Lancaster, PA, 1993.)

Rockstad, H., T. Kenny, J. Reynolds, W. Kaiser, T. VanZandt, and T. Gabrielson, "A miniature high-resolution accelerometer utilizing electron tunneling," ASME DSC **40**, 41-52, 1992.

INVITED PAPERS:

Gabrielson, T. B., "Sensor physics: signals and noise," Office of Naval Research Physical Acoustics Summer School, 16-23 June 2000, Asilomar, CA.

Gabrielson, T. B., "Exploitation and modification of off-the-shelf MEMS sensors for machinery diagnostics," ASME Energy Sources Technology Conference (ETCE 99), Special Session on MEMS for Non-Destructive Evaluation, 1-3 February 1999, Houston, TX.

Gabrielson, T. B., "Sensor physics signals and noise," Office of Naval Research Physical Acoustics Summer School, 14-21 June 1998, Asilomar, CA.

Gabrielson, T. B., "Sensor engineering on the microscale," 16th International Congress on Acoustics, 20-26 June 1998, Seattle, WA.

Gabrielson, T. B., "Limitations of the $g_h d_h$ figure-of-merit," ONR Materials Workshop, Penn State, May 1997.

Gabrielson, T. B., "Self noise: the relationship between a transducer and its electronics," Acoustical Society of America, Special Session on Sensor Self Noise, State College, PA, June, 1997.

Gabrielson, T. B., "Sensor physics: signals and noise," Office of Naval Research Physical Acoustics Summer School, 21-28 June 1996, Asilomar, CA.

Gabrielson, T. B., "Mechanical engineering in micro-dimensions and beyond," American Vacuum Society Micromachining Workshop II, Anaheim, CA, September, 1995.

Gabrielson, T. B., "Modeling and measuring self-noise in velocity and acceleration sensors," Acoustic Velocity Sensor Focused Workshop, Mystic, CT, September, 1995.

Gabrielson, T. B., "Fundamental noise limits for miniature acoustic and vibration sensors," ASME Winter Annual Meeting 93-WA/NCA-8, New Orleans, Nov/Dec 1993.

Gabrielson, T. B., "Fundamental noise processes in micromachined acoustic and vibration sensors," Acoustical Society of America, Special Session on Micromachined Sensors, New Orleans, November 1992.

CONTRIBUTED PAPERS:

Gabrielson, T. B., "Inexpensive technique for accurate determination of elastic properties by free decay in circular plates," J. Acoust. Soc. Am. **106**, 2231, 1999.

Gabrielson, T. B., "Application of differential-capacitance accelerometers to underwater acoustic velocity sensing," J. Acoust. Soc. Am. **106**, 2256, 1999.

Gabrielson, T. B., H. K. Rockstad, and T. K. Tang, "A miniature, high-sensitivity, electron-tunneling accelerometer," Acoustical Society of America, Special Session on Acoustical Measurement and Instrumentation, Austin, TX, November 1994.

Gardner, D. L., T. B. Gabrielson, and S. L. Garrett, "A simple sensor for direct measurement of particle velocity," J. Acoust. Soc. Am. **93**, 2323, 1993.

Gabrielson, T. B., "Mechanical-thermal noise in microminiature acoustic sensors," Third International Workshop on Transducers for Sonics and Ultrasonics, Orlando, 6-7 May 1992.

PATENTS:

Gabrielson, T. B., J. F. McEachern, and G. Lauchle, "Underwater acoustic intensity probe," U.S. Patent No. 5,392,258, Feb. 21, 1995.

Gabrielson, T. B., "Apparatus and method for calibration of sensing transducers," U.S. Patent No. 5,644,067, July 1, 1997.